

DNA sequence tag

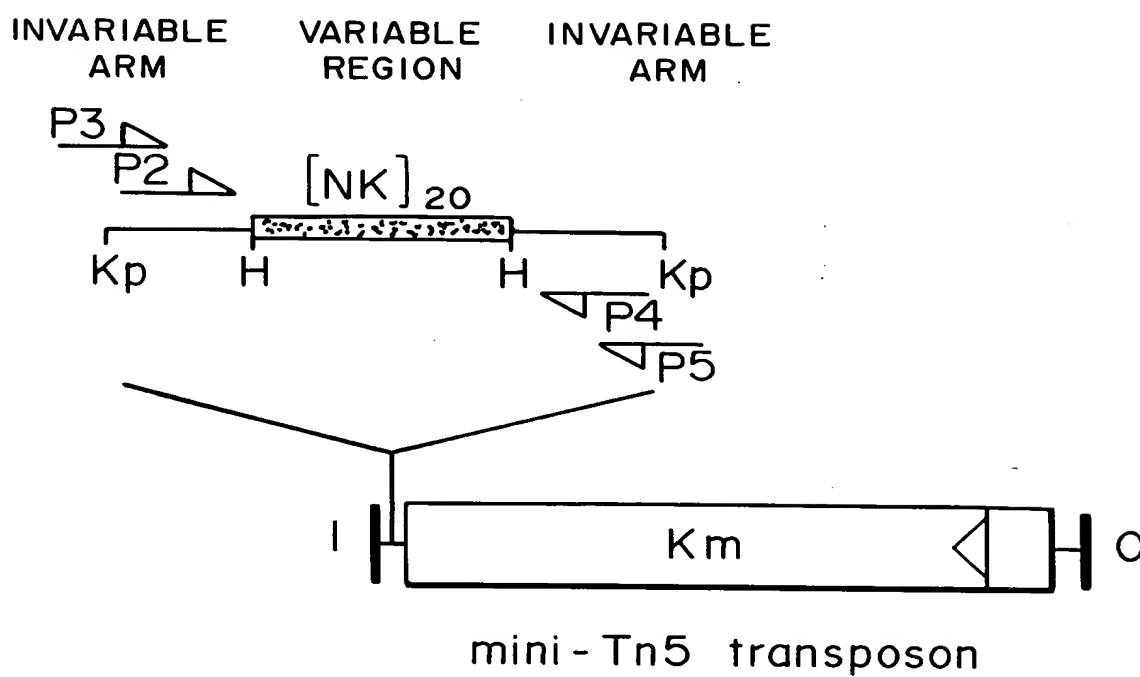


FIG. 1A

MICROTITRE DISH WITH TAGGED
TRANSPOSON MUTANTS

FIG. 1B

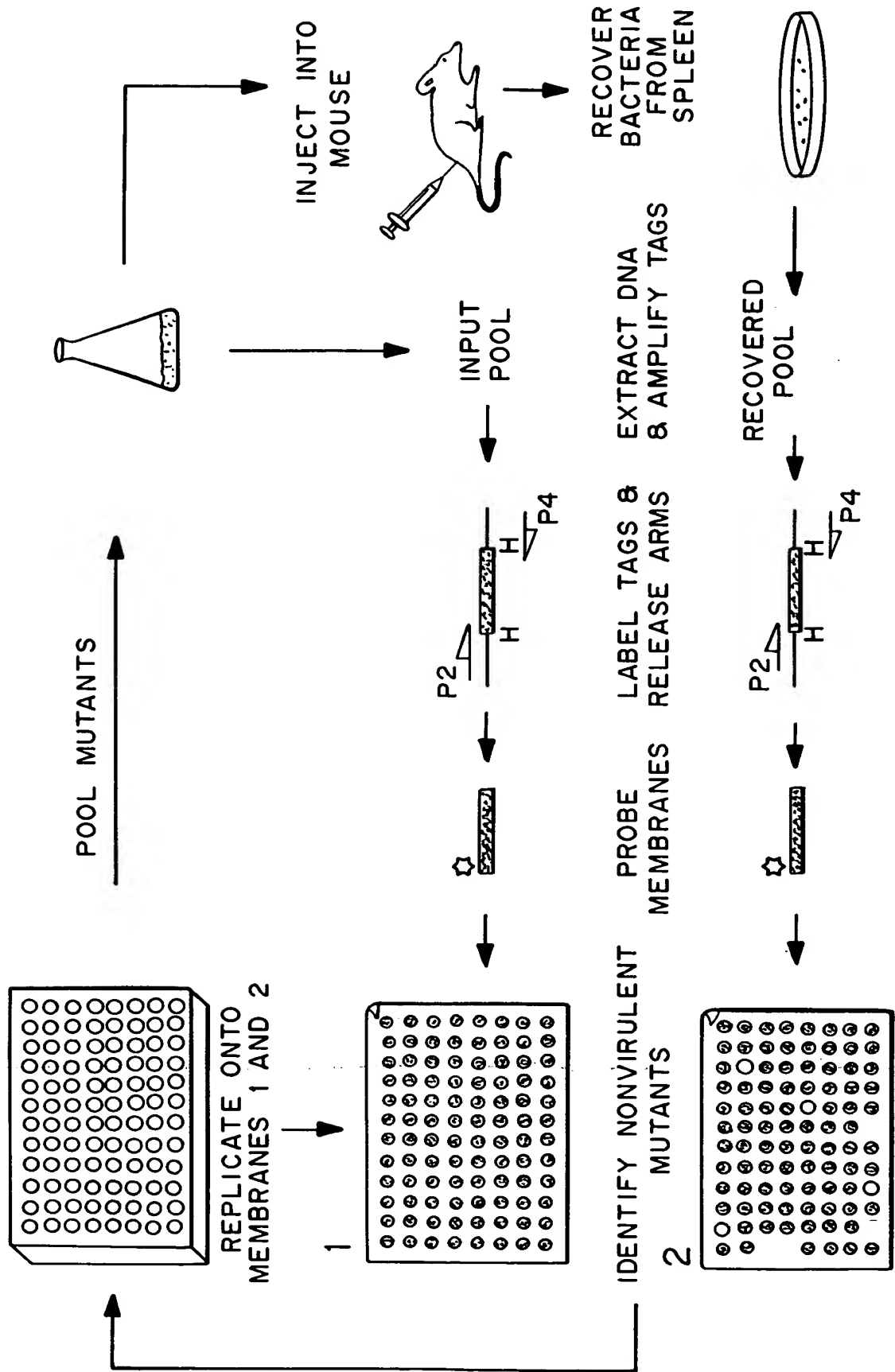


FIG. 2

1 2 3 4 5 6 7 8 9 10 11 12

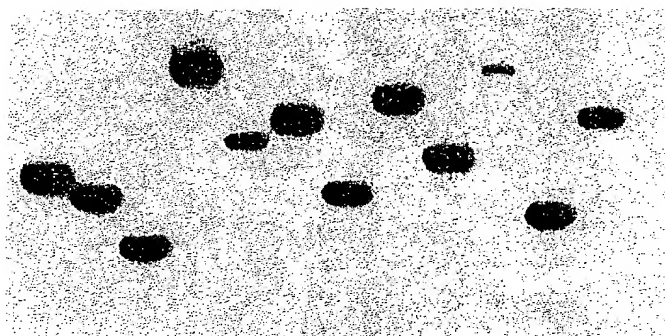
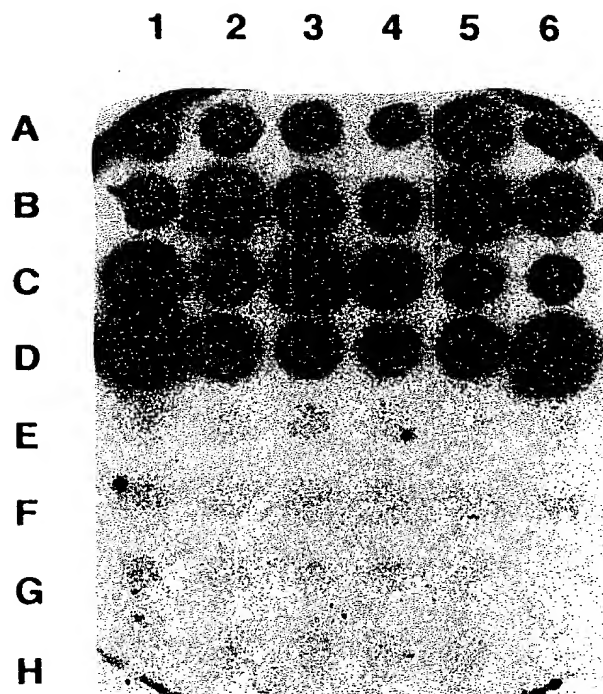


FIG.3



09714602-11600
009111-2094760

FIG.4A

Inoculum pattern

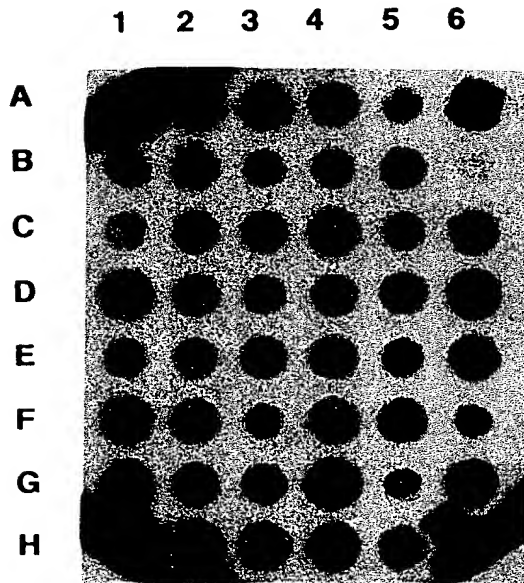


FIG.4B

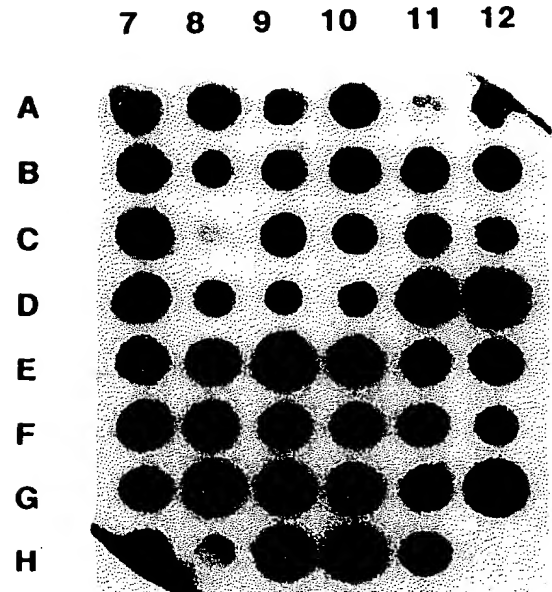


FIG.4C

Spleen pattern

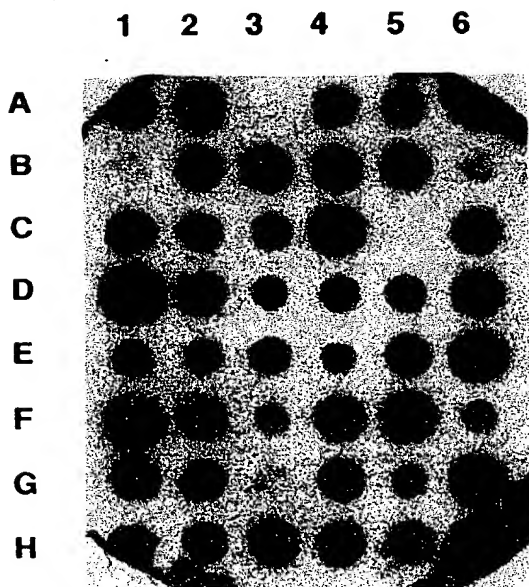
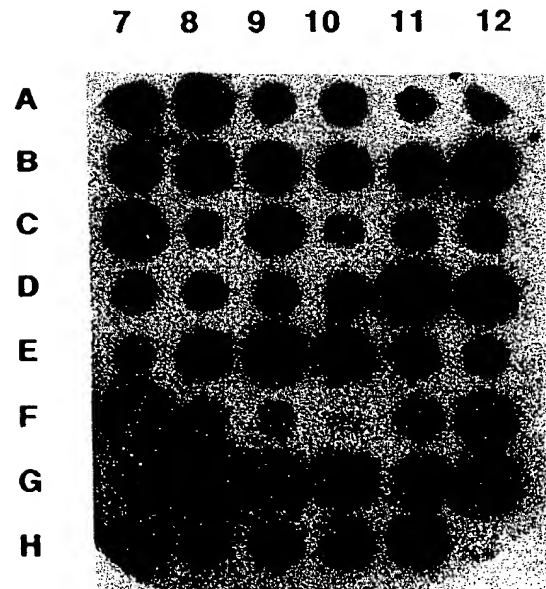


FIG.4D



Name: mpcc2 1

J05534 Escherichia coli ATP-dependent clp
protease proteolytic component (clpP) gene
complete cds. Length = 1236

Minus Strand HSPs:

Score = 453 (125.2 bits), Expect = 4.3e-28, P = 4.3e-28
Identities = 113/141 (80%), Positives = 113/141 (80%),
Strand = Minus

Query is our Salmonella sequence ———→
Query:359 CCACCAGCCGCTGGGTACCAAGGCCAGGCCAGGATATTGA 318
||||| | | || ||||||||||||||| ||||| |
Sbjct:785 CCAACCGTTGGCGGCTACCAAGGCCAGGCCAGGATATCGA 826
c/pP gene ———→
Query:317 AATCAGCCCGCGAAATTTGAAAGTAAAGGCCGCGATGAA 276
||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct:827 AATCATGCCCGTGAAATCTGAAAGTTAAAGGCCGCGATGAA 867

Query:275 TGAACCTATGRMKYKMMATACGGTCACTCTTGA 240
||||||| | ||||| | | |
Sbjct:868 TGAACCTATGGCGCTTCATACGGTCAATCATTAGA 904

Query:239 GCAGATTGAASGTGATACTGA 219
||||||| ||||| | |
Sbjct:905 ACAGATTGAACGTGATACCGA 925

FIG. 5A

Continuation of 09/201,945
RPM5101Cox(3)

FIG. 5B

Score = 231 (63.8 bits), Expect = 4.0e-24,
Poisson P(2) = 4.03-24
I tities = 55/66 (83%), Positives = 55/66 (83%),
Strand = Minus
Query:194 TGAAGCGGTAGAGTACGGTTTGGTTGACTCAATTGACCCA 154
||||| ||| ||||| |||| ||| ||| |||||
Sbjct:950 TGAAGCGGTGGAATACGGTCTGGTCGATTCTGACCCA 990

Query:153 TCGTAATTGATGCCCTGG 135
||||| ||||| |
Sbjct:991 TCGTAATTGATGCCAGAG 1009

Query:134 ACGCAA 129
|||||
Sbjct:1010 GCGCAA 1015

FIG. 5C

>ECCLPXGNA Z23278 E.coli ClpX gene, complete CDS

Length = 1945

Minus Strand HSPs:

Score = 364 (100.6 bits), Expect = 1.6×10^{-20} , P = 1.6×10^{-20}
Identities = 88/107 (82%), Positives = 88/107 (82%),

Strand = Minus

Query:325 GATATTGAAATTCACGCCCGCGAAATTTGAAAGTAAAGGG 285

||||| ||||||| ||||||| ||||||| ||||||| |||||||

Sbjct: 1 GATATCGAAATTCATGCCCGTGAAATTCGAAAGTTAAAGGG 41

Query:284 CGCATGAATGAAC TTATG 266

||||| ||||||| |||||||

Sbjct: 42 CGCATGAATGAAC TTATG 60

Query:265 RMKYKMMATACGGGTCAATCTCTTGAGCAGATTGAASGTGATACTGA 219

||||| ||||||| || || ||||||| ||||||| ||

Sbjct: 61 GCGCTTCATACGGGTCAATCATTAGAACAGATTGAACGTGATACCGA 107

FIG. 5D

Score = 231 (63.8 bits), Expect = 6.8e-24,
Poisson P(2) = 6.8e-24
Identities = 55/66 (83%), Positives = 55/66 (83%),
Strand = Minus

Query:194 TGAAGCGGTAGAGTACGGTTGGTTGACTCAATTTGACCCA 154
||||| || ||||| || || ||| |||||
Sbjct:132 TGAAGCGGTGGAATACGGTCTGGTCGATTCTGACCCA 172

Query:153 TCGTAATTGATGCCCTGG 135
||||| ||||| |
Sbjct:173 TCGTAATTGATGCCCAGAG 191

Fetch → Gb_ba:Ecoc1ppa
Query:134 ACGCAA 129 - OK then type J Biol Chem 265, 12536
||||| (1990)
Sbjct:192 GCGCAA 197

A) new virulence factors with similarity to sequenced genes:

1. p1F10

similarity to *clpP* (*E. coli*)
(Figure 5)

2. p2D6

similarity to *lcrD* (*Yersinia* spp.)
sequence p2D6_1_I

GGTCTTAATGTACGGGCATGGTCTGCATCGATAACTCCGGCACGCAAATCG
CCATCGATACTCATTTGTTTGGCTGGCATCCCATCAAGCGAGAAACGTGCG
CTAACTTCCGCCACCCTCTCGATACCTTTTGTAATGACAATAAATTGCACG
ATAGTAATGATGGTAAATACGACCAACCCAACGGTGAGATTTCTCTCTACG
ACAACTTACCGAAAGCATCCACAAATATTACCGGCATTATGTTGTAAACAG
TACCCAGCCGTGATGTGCTGATTGGGGAGTTAACAACCGATTTAT

3. s4C3

probably same gene as p2D6, but different region
similarity to *S. typhimurium invA* and *Yersinia*
spp. lcrD
sequence s4C3_1_U

GCGCGGACGCTAGTGTGGTGGGTGACAGCCAGACGTTACCGAACGGGATGG
GGCAGATCTGTTGGCTTACAAAAGACATGGCCCATAAGGCGCAAGGTTTTC
GGACTGGACGTTTTTCGCGGGCAGACAACGTATCTCTGTCTTATTAAATGT
GTCCTGCTTCGGCATATGTATCGAACCCTCGGAGCAAAGTCGTTTGGGCGC
AGAATTAGTACGTTTGGGTGCGTTGCTGTTATTCTTGGGCTCGGAAAAAG
AGTGCCAGCGTGAAGGAGTGGGATTTGGCAGACTGGCCGCCTAAT

sequence s4C3_1_R

CACTATAGGGAAAGCTTGCATGCCTGCAGGTGCGACTCTAGAGGATCTACTA
GTCATATGGATTGCACTTGTGTATAAGAGTCAGGATTAGAGGACATGCGCC
GGGAACCATACTATCTTTTTCCGGTGCTTCGACGCCATTTGCGGAAACCAC
AGACTTTTTGCGGCGAATGAGGATAATTGGCAATGCTAACAACGCTGAAAA
GAAAGCGAGAGTGATAAAAGGAAAGCCAGGAATTAAAGCGAGGAGCATTAA
AACCACAGCGGCTAATATGAGCGACTGAGGTTGTCTGGCAATTTG

Figure 6A

4. p3F4
similarity to *invG* (*S. typhimurium*)
sequence p3F4_1_U

TGCAGGCCGACTCTAGAGGATCCCCGGGTACCGGTAATTTCTTTAACCTCG
CATCCCCGGTGGATGAAAGGATATTCTGGCTGCGTAAGTAATGAATGAACCG
CCCAGTAGATAAAAATATTGAAAGTGATAACCTGATGTTTTTAATAACGATGC
AGGATATACATATAACATGCTGGCATCAAACCAGGTAAGCAAATCATATTG
TGCTGCCAGGTTATTCAAACATATCGACCGGTGGTCCAGGCGGGAATTTTTC
CACTAAATGTAGGTGGGATCAATGGGCTAATTGGTATAGGCGGAT

5. p7G2
similarity to *yscC* (*Yersinia spp.*)
sequence p7G2_1_U

CCTGTGATTCCGGATGAAATAGCTTTTACGAAAGCTGTCAGACNTGCTGAA
GAATACGCTGCAAATGGTAAGCTTGTAACCTTTTGGGTATTGTTCCAACGCA
TGCTGAAACGGGTTATGGATATATTCGTCGCGGTGAGTTGATAGGAAATGA
CGCTTATGCAGTGGCTGAATTTGTGGAGAAACCGGATATCGATACCGCCCG
TGAATATTTCAAATCAGGGGAAATATTACTGGCCTAGCGGCGATGTTTTTA
TTTCGCGCAAAGCCCTTATTTAAACGAATTAACGTATCTATCACCCCCAA
ATTCATACAGCTTGTGAA

sequence p7G2_3_O

TTACTAAACAGGGCCCCGGACCATGTAAACACCACGCTTGCCAACACTAAA
AAACGATGCTTGCCGTAAAAAAATTGAACGTTATTTACTTAATACGCCTAT
TTTATTTACATTATGCACGGACAGAGGGTGAGGATTAAATGGATAATATTG
ATAATAAGTATACTCCACAGCTATGTAAAATTTTGGGGGCTATATCGGATT
TGGTTGTTTTTAATTTAGCCTTATGGCTTTCCTAGGATGTGTCTATTTTTT
TTTGTGGTCAAGCACAGAGATTTATTCCCCAACCACC

sequence p7G2_1_I

TTTCCTTGCCGTGACAGTCCGGGATGCGAGGTTAACGAAATTACCGGCACC
AAAGCTGTGGAGGTGAGCGGTGTCCCCAGCTGCCTGACTCGTATTAGTCAA
TTAGCTTCAGTGCTGGATAATGCGTTAATCAAACGAAAAGACAGTGCGGTG
AGTGTAAGTATATACACGCTTAAGTATGCCACTGCGATGGATACCCAGTAC
CATTATCGCGATCAGTCCGTCGTGGTTCCAGGGGTGCGCTAGTGTATTGCG
TGAGATGAGTAACACCAGCGTCCCGACGTCATCGACGAACAATGG

Figure 6B

6. p9B7

similarity to *fliQ*, *invX* (*E. coli*)

sequence p9B7_1_1

CATGAGTAACCTACCCAACCTGTAATCTTTACCAATATGCATCATAATCTTC
TGCTGGTAAATGATTGGTAATATCGGAAAGGTAAGTGACATAAGCACGCCA
TTACGTAAAAGTGCGGCCCCCTAAACTGCCACTTTTTTAATAAGGGAAGTAAT
AAAGAAAGGCTCAATGGTCGAATAAAAGCCACAGCCAATGCAATAAGCCAC
TCATTTACCTGTTGTGCCATTCAACCATGCTCTCCAATTTCGTAACATTATC
TGCCGGGTATAATTCAACAGGATACCGCTAAGCCATGGGTTAG

sequence p9B7_3_0

ATTCCAGCCCCCGGGCCATCTAACCACCTATGAACAATCATCTTCTGGGTGG
ACAATCATTTGGTACCATCGGCCAGGCTTGTGCAATATGTATGTCATCACGT
AAAAGCGCGGCCCTTAATCTCCCCATTCTTCCTTAAGGGCAGTTATCACG
GCTGGCTCAATGGCCGGCTTAACAGCCACAG

7. s6F5

similarity to *yscU* (*Y. enterocolitica*)

sequence s6F5_1_0

GAGGCGCGTCTTCGGTTGAGGGTCGCCCTCCAGATCTTTATGCTCCTGTTT
TACGTCATCTTTACTCATTTTAAAGATCTTTTCTAATCTTATAATATTGAAA
AGAATAGTCCAGTATGCCAACGACGAAATAAAGAAACATCACCCCAACCCA
TAACCATTTTTTCAATGATGAAAGCACAGCACGCCACAGGCTACACCACA
GCCCCGAGGGGGCCGGAAAGTGCTGGGATCTTGATTAATGAAAAAGGCAAA
GGGAAGAGATAGGATGATGCATGCTGGTTGGAGGCAGATTATTCATCTTCG

Figure 6C

0974602-11600

B) new sequences without similarity to entries in DNA or protein database:

1. s4D10

sequence s4D10_1_U

AGTTGCCGTATTTATTAAATATTCACCTCAGGTCAATATGGAGGTCTTCCC
GGCTAAAAATCATTGCTTTACTAGAGATATCACTCCCTGGGTTGCAATACA
GTACGATTAGTTATCTTGATGCAGCCTGCTGATTTTCAGAATGGCAGCTGAC
GTACCCGCGAGACAAACATTCTGGATTATGGACGTTATCAACGCCAATATA
GGGAAGGTGGTGAAGTGGTTGATGAAATACCCCTATCCCTTGTCATGTTATC
GCTGACAGGACTGTTATCAGGAGCGGGCATCCTCGATCGGCT

sequence s4D10_1_R

CAAGAGACAGATCCAACCTCGGGCCGATCGCCATAACGCCAGCAGTTTGAAA
GATGAAAGCCCAGCTTATCCAGCCATTCCGGTACAGCGTAACGAGCAGGTT
GCCAGAAATAACGATAAAGTTGCAACACCTCGGGATCAGGTCGGCTCAAAA
ACGGGGTCTCAGGCAAAAATAGCCGATCAGGATGCCCACTCCTAATAACAG
TCCTGTCAACGATAACATCAACGGATAAGGGTATTTTCATCAACCACTTCAC
CACCTTCCCTTTATTGGCGTTGGATAACGTCCATAATCCAGA

2. s4H10

sequence s4H10_1_U

AGGGCTTTATTGATTCCATTTTTTACACTGATGAATGTTCCGTTGCGCTGCC
CGGATTACAGCCGGATCCTCTAGAGTCGACCTGCAGAACCGAGCCAGGAGC
AAATTAATTTTTTTTGGGCAATTGCTGAAAGATGAAGCATCCACCAGTAACG
CCAGTGCTTTATTACCGCAGGTTATGTTGACCAGACAAATAGATTATATGC
AGTTAACGGTAGGCGTCGATTATCTTGTCAGAATATCAGGCGCAGCATCGC
AAGCGCTTAATAAGCTGGGTAACATGGCATGAAGGGGCAACCC

sequence s4H10_1_R

CACTATAGGGAAAGCTTGTCATGCCTGCAGGTGCACTCTAGAGGATCTACTA
GTCATATGGATTCCTAGGCGGCCAGATCTGATCAAGAGACAGATCCAACCTC
GGGCCGATCGCCATAACGCCAGCAGTTTGAAAGATGAAAGCCCAGCTTATC
CAGCCATTCCGGTACAGCGTAACGAGCAGGTTGCCAGAAATAACGATAAAG
TTGCAACACCTCGGGATCAGGTCGGCTCAAAAACGGGGTCTCAGGCAAAAA
TAGCCGATCAGGATGCCCACTCCTAATAACAGTCCTGTCAACG

Figure 6D

3. p4G5
sequence p4G5_1_0

CCCCCCCCCTTCTCCTGGCTTACACAGCCCCAGACCGGCGCTGGAAAAGGC
CATTTCCCGCCATACAGGAGGCCAGCAACATATTTTCACGCGCCGCCAGATC
GTGGCCGTAACCCACGGCTTTCGGCAGCGATTTGCCAATCATCGCTATCGC
GCCAATCGCCAGGCTGTTCGGTAAACGGCGTGGCGTTGAGCGCGCTGTAGGC
CTCAATCGCATGCGTCAACGCATCGATAACCGGTCATCGCCGTCACGTTTGG
CGGAACGCCTTCGGTCACGGAAGCATCAAGAATCGCCACGTCCGGC

sequence p4G5_1_U

CGCGAACGTGCGCCGCAACTGCTTGTGGACGGTGAATTGCAGTTTGAACGCC
GCTTTCGTGCCGGAGGTCGCCGCGCAAAAAGCGCCTGACAGCCCGCTGCAA
GGCCGCGCCAACGTGATGATTTTCCCGTCGCTGGAGGCGGGCAATATTGGC
TACAAAATCACTCAGCGTCTGGGAGGCTATCGCGCTGTTGGGCCGCTAATT
CAGGGGCTTGGCGCGCCGCTTCACGACCTCTCCCGAGGCTGTAGCGTGCAG
GAAATTATCGAACTGCGGTTGGTGAGAAAACCAA

4. p7A3
sequence p7A3_1_U

CGCCCTAGCATGCCTGGCGTTGTCCGGTTATTGCTCGTCAAGCGAACAGAT
GCAAAAGGTGAGAGCGACTCTCGAATCATGGGGGGTCATGTATCGGGATGG
TGTAATCTGTGATGACTTATTGGTACGAGAAGTGCAGGATGTTTTTGATAA
AAATGGGTTACCCGCATGCTGAAGTATCCAGCGAAGGGCCGGGGAGCGTGT
TAATTCATGATGATATACAAATGGATCAGCAATGGCGCAAGGTTCAACCAT
TACTTGCAGATATTCCCGGGTTATTGCACTGGCAGATTAGTCACTCTC

sequence p7A3_1_I

CCCTTCCCAGGCTCGACAGGTACACAGCCAGCCACTGGTGCAGGCAGTTAC
TTGCTTTCATCATGGGAAGGAGCAATATCCTGATATATTAAAGAAAGAGCG
GGATCCCCCTTTCTTTACTGCTGCTAACGTTTCTTGCAAAATGCGTTGATGA
GATTCATCCAGCACACCACTGATAACAAAAGAGCGCCGCATTGGCGTAACA
TTGACAAGCCCCACTAAACCGCTCTCTATTATCGCAGAAATAATATCATCC
CCCTGAGACTGATGAGAGTGACTATTCTGCCAGCGCAAATAACCC

5. p10E11
sequence p10E11_1

ATACCGAGTATTAAGCGGCTGTGTAACATCGTCATCCAACAACATACGCAG
CGAGCCGCCACGCCGGA AAAAACCGCATCGTGTTCATGTGCCTGTTGTAGGGT
CGGGTCTTTTTTTCATGAGTACGTTTTCTGCGCTATCATACTGGAAATTTCC
CCCCACTTACTGATAAGCCCTGTCAGTTGGGTAAAGACAGAGTTAAGCTCC
TGAGACATTTTTTGGAAATGGTTATCTTTCCCCGACTCATAAAATCGGTATT
CCCGCTGGGGGCAATATCCAAAGACGCTTTGGTTCGCCCCGTAGGGCACC

Figure 6E

sequence p10E11_U

GCCGTATGCCTGCAGTTGCCCGGTTATTGCTCGTCAAGCGAACCGATGCCA
AAGGTGAGAGCGACTCTCGAATCATGGGGGGTTCATGTATCGGGATGGTGTA
ATCTGTGATGACTTATTGGTACGAGAAGTGCAGGATGTTTTGGTAAAAATG
GGTTACCCCCATGCTGAAGTATCCAGCGAAGGGGCGGGGAGCGTGTTAATT
CACGATGATATTCAAATGGGTTCAGCAATGGGGCAAGGTTCAACCCCCACTT
GCAGATATTCCCCCCCCCTATTGGACTGGCAGATTAGTCACTCTCA

6. s4B9

sequence s4B9_1_O

GGGCGACCTGCCCGCGGCGCAACTTTCCCCGAAGCGTTTTCCATTTTCCTTG
TTCTTAAATGACCTGGAAAGCTTACCTAAGCCTTGTCTTGCCTATGTGACA
ATACTGCTTGGAGAACACCCGGACGTCCATGATTATGCTATACAGATCACA
GCGGATGGGGGATGGTGAATCGGTTATTATACCACAAGTCGCAGCTCTGAG
CTTATTGCTATTGAGATAGAAAAACACCCCGCTTCAACTTGGATTTTGAAT
AATGTAATACGCAATCACCATACTATATTCGGGTGGCGTATAA

sequence s4B9_1_R

TTCGAGCTGGGGCACCGCTAATATCTTTAACCTCGCATCCCGGTGATGAAA
GGATATTCTGGCTGCGTAAGTAATGAATGAACCGCCAGCAGATAAAATAT
TGACAGTGATAACCCGATGTTTTTTTAACGATGCAGGCTATACATATAACA
TAGCTGGCCACCAACACAGCTGAAGTAAATCATATTGTTGCTGCCAGGCTA
CTTCACACTATTGTCCGGCGGGCCAGCGGGGATTTTCCCCCTAAATCTCGC
TGGTTCTCAAA

7. p4F8

sequence p4F8_1_I

AGTCTACGATTTTCGCTATATCTTCTCTTAATCATGGCCGCCATTTGTGGAT
GCGATTTTAAATATCCGGGCGATCTTTCATTAAAAAATAAAGATTCCCCA
TGACTTCACAGATAAAGGTATCGGTATTTTGAGTGATACGTAACAATTTCGT
TCTCTTCGTGTGGGTCCATGATGCGAAGAATAATGGTGGCATCATTTTCAT
GAGGATTATGAACCCGAAATCTTTCTCTTTGCGATGCGCAGGCTAACTCTT
TCAACTCAAAAAAATCTCTGTAAGCCGCTCTCGTGTGGGGGCGC

8. p7B8

sequence p7B8_1_O

GCGCCCCTTTAATTGGTTGAGGCGGCTGGTATTCTTGTAAGGGTAATACTA
GCGAGACCCAGGTTCCACCCCCGGGGACACTTTTTAGTGTCAGATTACCGC
CCATCATTTTAGCCAGGCTTGACGCAATAGTCAGTCCAATTCCTGTACCTT
GCGAATTTGTGTCTGCTTGATAAAAAGCAGAAAAGATTTGAGACTGCTGCT
GTTTTTCAATCCCCCACCCTATCGCTAACCAGAAATATTAATTGTTTCCT
CACCAAGATTGAGCGCCAGACGTATCCCTCCCCCCTCGGGAAAT

Figure 6F

9. p8G12
sequence p8G12_1_I

GGATAAGATCCCGGATAAGTATGTCAGGCTCGTATGCACAACAGGCATTAT
AAACCTCTAGACCATTTTTTAACATGCTCTACTATTTTAAAATGAGGCCAGG
GTAATAAGGCATTCATAATGCCGTTAATGATGATTTTCATGATCGTCTACTA
ATAAGATCTTATATTCTTTTCATTTGGCTGCCCTCGCGAAAATTAAGATAAT
ATTAAGTAATGGTGTAGGTGTGGAGATCATAACGTATTTTCTGGCGTAAGT
CGTTAGTTCCTCCAGCGCGATGATTTTCCCCATTTTACGCGAT

10. p9G4
sequence p9G4_1_O

TTCCATATTGCTCGTCCGGGGAGCGTGTTAATTCTTGATGATATACCAATG
GATCTGCAATGGCGCAAGGTTCAACCATTAATTGGAGATATCCCCGGGTTA
TTGTACTGGGAGATTAGTCACTCTCATCAGTCTCAGGGGGGTGATGTTATT
TCTGGGATAATAGAGCAACGGCGTTAGCAGGGGTTCGGTCAGTAGTCACGGC
CAACTTCGGTGCACCTTTTGCCTATCACTGGGGTATCATAACTGAATCTCAT
CCCCCCCCACTTTGGTAATCACAC

sequence p9G4_1_U

AATTCTTTTACCTCCATAAGCTGCGTGGCATAGCGATACAGAGTATTAAGC
GGGTGTGTTACATCGTCATCCAACAACATACGCAGCGAGCCGCCACGCCGG
AAAAACCGCATCGTGTCTATGTGCCTGTTGTAGGGTTCGGGTCTTTTTTTCAT
GAGTACGTGTTCTGCGCTATCATACTGGAAATTTCCCCCACTTACTGATA
AGCCCTGTGAGTTGGGTAAGGACAGCGTTAAGCTCCTGAGACATTTTTTGA
GTTGTTATCTGCCCCCGACTCATAAGATCGGGTATTCCGCGGTGG

11. p9B6
sequence p9B6_1

ATATCCCTAATGCTTTTTCCTTAAAATAAATACCACGGAAGGATACTGGCCA
CCTAGCCAAATTTAGAAAGCAATGAACATCCGGTTTATTCTGAAAACGAT
TACTCCGGCGCACGTTGTTCTGGCGTTACCTGAGCCAGCAAACGATATAAT
GGGGTGGTGACCCGCATACCGGTCATTGGCATCCCATCCACACCGGAGGGA
GTAAAACTCATTAGGCCATAGGTAATATCATTAAGACGCTCTAATAAATGA
GGGTGGGGGGGCCAAACTACCACTCCAGTATGTATTGAGTCA

Figure 6G

12. p6G5
sequence p6G5_2_I

CCCATGGGCGCAATTTGTTGCGCAGCGTTTACCCGACCATCGCGTTTATGA
GCTGTAATTCATGGGGGGTAAAAACGGGCGTGACGACCCCAACGGAAGATA
AGGCCGGGCTTAAACAGGAGATTATTGCTAATGCGCAGCGCAAAGTGTTGC
TGGCGGACAGCAGTAAGTATGGCGCGCATTGCTCTTTAATGTGGTGCCGC
TTGAGCGCTTTAATGACGTGATTACCGACGTCAATCTGCCGCCGTCAGCGC
AGGTTGAACTGAAAGGGCGCGCTTTTTGCGCTAACG

00911 2094160

Figure 6H

FIG. 7B

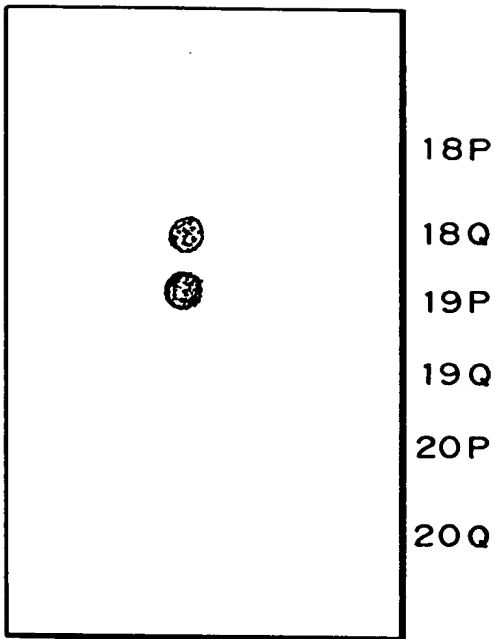
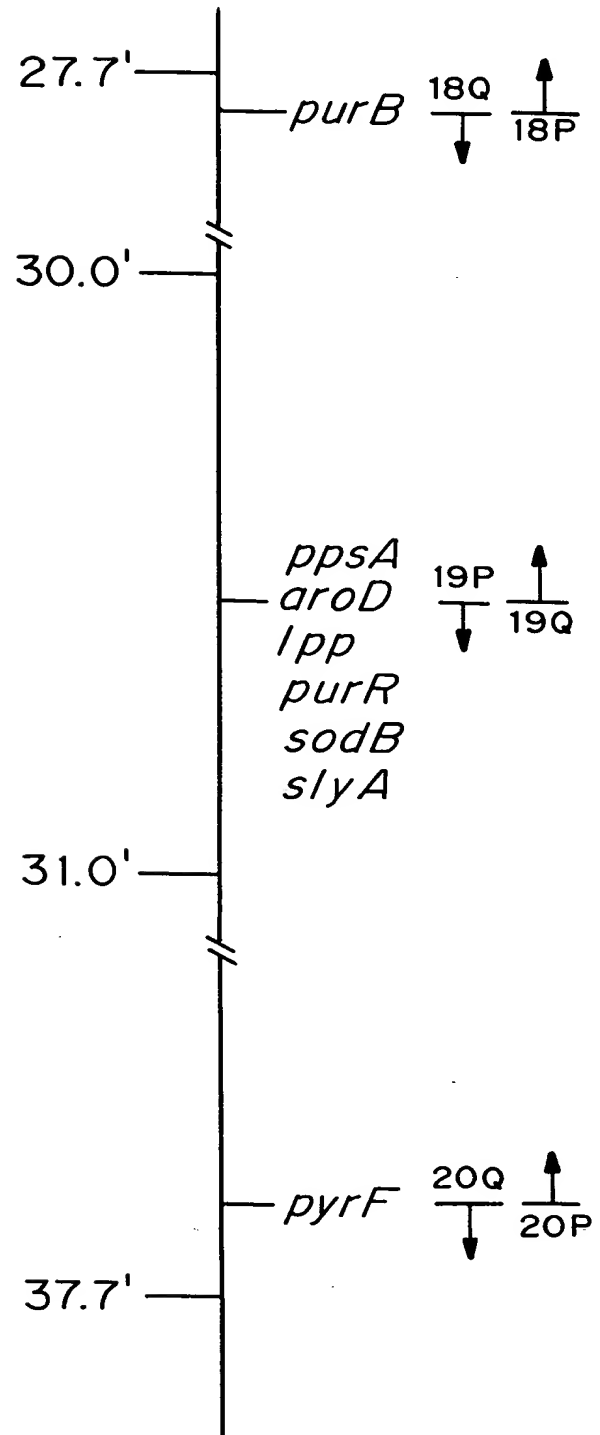


FIG. 7A

009111-2094160

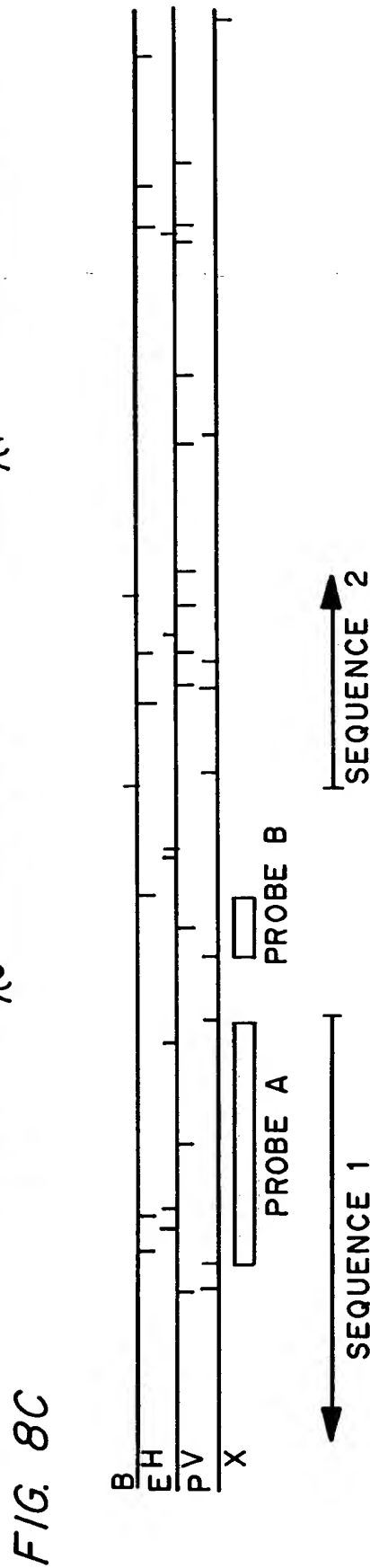
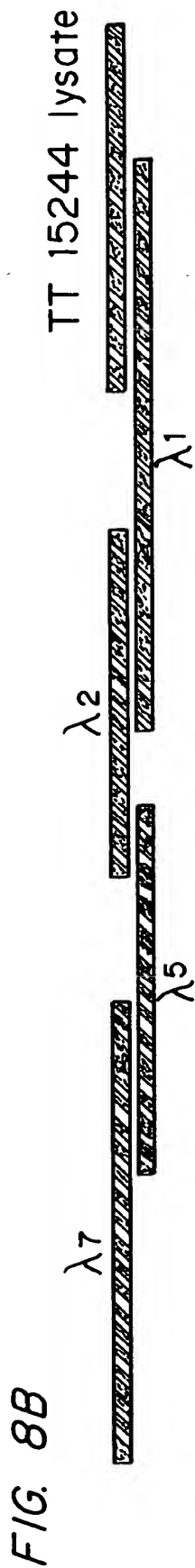
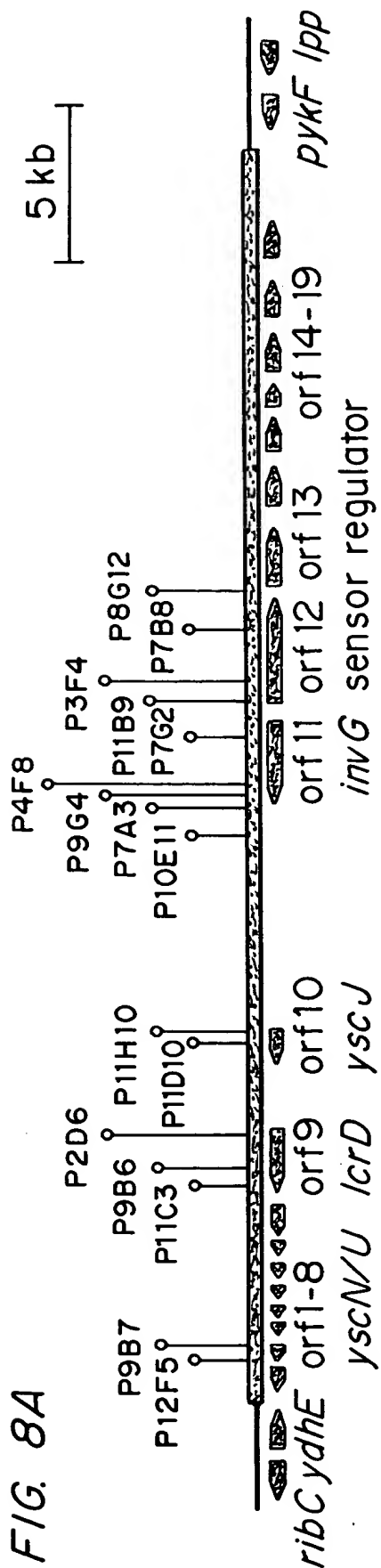


FIG. 9

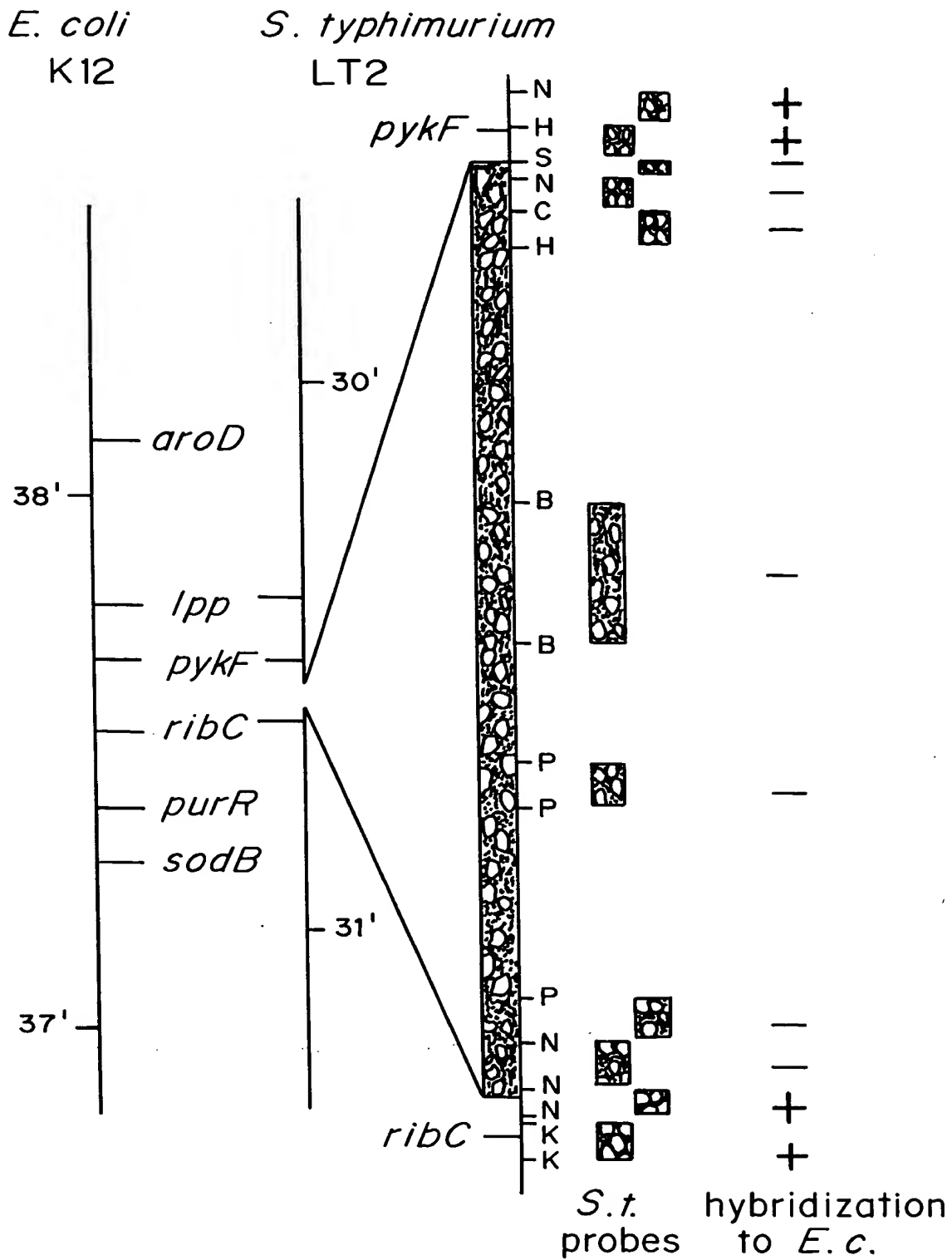
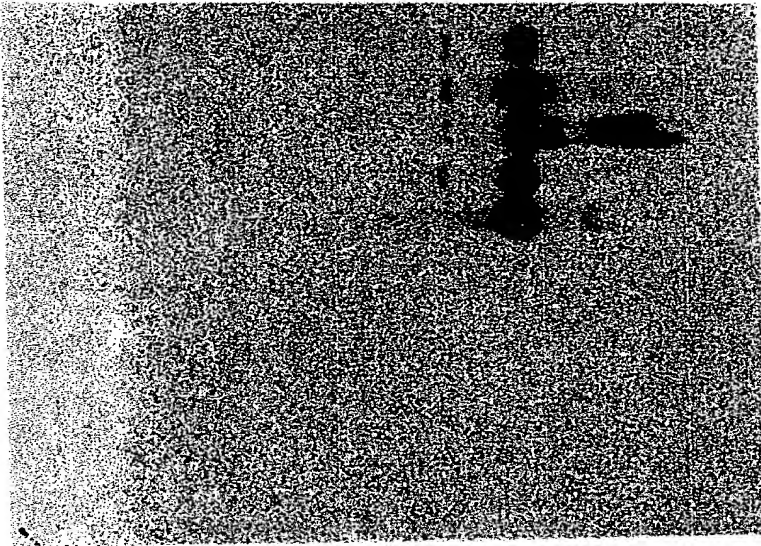


FIG. 10A

*Pst*I

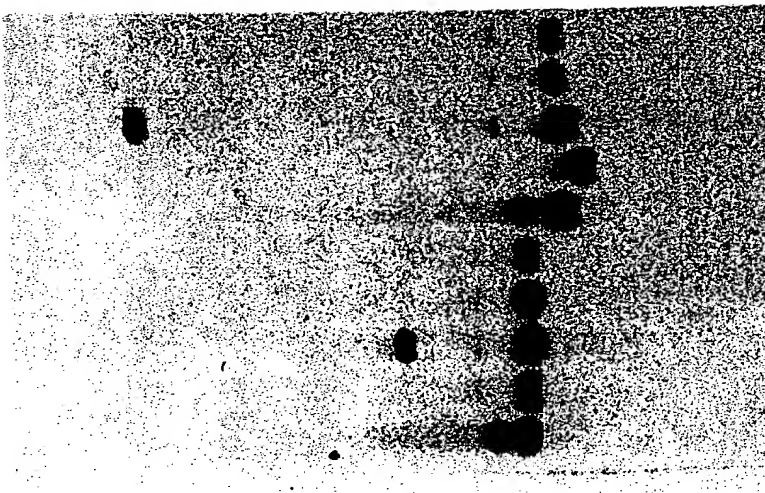


S. typhimurium
S. aberdeen
S. gallinarum
S. cubana
S. typhi
EPEC
EHEC
Y. pestis
S. flexneri
V. cholera
S. aureus

FIG. 10B

*Hind*III

*Eco*RV



S. typhimurium
S. aberdeen
S. gallinarum
S. cubana
S. typhi
S. typhimurium
S. aberdeen
S. gallinarum
S. cubana
S. typhi

a C S - R - A S I I L P E Y H G A A C Q A
b A V N G R R L S C Q N I T A Q H A K R
c Q L T V G V D Y L A R I S R R S M P S A

Figure 11A

Figure 11B

367

420

൩

د

7

421

480

၈၀

D

U

481

540

π

۵

2 3

Figure 11C

009

a L K E C - V R W R G V I N G K S D H C A
b - R N A E S D G G A - L M A K V T I A L
c E G M L S Q M E G R D - W Q K - P L R Y

099

a T D L - - G K - R F S E L S C R I Y K I
b P T Y D E G S N A S P S V A V F I K Y
c R L M M R E V T L L R A Q L P Y L - N I

720

a	F	T	S	G	Q	Y	G	G	L	S	G	K	N	-	R	F	N	R	D	V
b	S	P	Q	V	N	M	E	A	F	R	V	K	I	K	D	L	I	E	M	S
c	H	L	R	S	I	W	R	P	F	G	-	K	L	K	I	-	-	R	C	Q

Figure 11D

Figure 11F

Tn insertion P11H11
↓

ACTGTTATTAGGAGTGGCATCCTGATCGGCTATTTTTCCTGAGACGCCGTTTGTGAGC 960
-----+-----+-----+-----+-----+-----+-----+
TGACAATAATCCTCACCCTAGGACTAGCCGATATAAACAAGGACTCTGCGGCAAAAACCTCG

901

a T V I R S G H P D R L F L P E T P F L S
b L L L G V G I L I G Y F C L R R R F - A
c C Y - E W A S - S A I F A - D A V F E P

CGACCTGATCCCGAGGTGTGCAACTTTATCGTTATTTCTGGCAACCTGCTCGTTACGCT 1020
-----+-----+-----+-----+-----+-----+-----+
GCTGGACTAGGCTCCACAACGTTGAAATAGCAATAAAGACCGTTGGACGAGCAATGCCGA

961

a R P D P E V L Q L Y R Y F W Q P A R Y A
b D L I P R C C N F I V I S G N L L V T L
c T - S R G V A T L S L F L A T C S L R C
Tn insertion P11D10
↓

GTACCGAATGGCTGGATAAGCTGGGCTTTTCATCTTCAAACTGCTGGCGTTATGGCGATC 1080
-----+-----+-----+-----+-----+-----+-----+
CATGGCCTTACCGACCTATTTCGACCCGAAAGTAGAAGTTTGACGACCGCAATACCGCTAG

1021

a V P E W L D K L G F H L Q T A G V M A I
b Y R N G W I S W A F I F K L L A L W R S
c T G M A G - A G L S S S N C W R Y G D R

Figure 11F

1081

ന മ ഉ

1141

சுரு

1201

ਥੁ ਧੁ ਧੁ

Figure 11G

1261

π

9

U

1321

॥

د

U

1381

π

2

U

Figure 11H

1441 TTAGTATTATTACCCCTCCGACGGTATACTTTGGCCGAAGACTTCTCTTACCGAGATT 1500
-----+-----+-----+-----+-----+-----+-----+-----+
AATCATAATAATGGGGAGCGTCGCATATGAACCGGCTTCTGAAGAGAATGGCTCTAA
a L V L L P P P Q R I L W P K T S L T E I
b - Y Y Y P L R S V Y F G R R L L L P R L
c S I I T P S A A Y T L A E D F S Y R D Y

1501 ATCTTCATGGAGCATTTGCTATGAGTTTACTTCACTTCTCTGACGGAATTAACCATA 1560
-----+-----+-----+-----+-----+-----+-----+-----+
TAGAAGTACCTCGTAAACGATACTCAAAATGAAGTGAAGGAGACTGCCCTTTAATTGGTAT
a I F M E H L L - V L L H F L - R K L T I
b S S W S I C Y E F Y F T S S D G N - P -
c L H G A F A M S F T S L P L T E I N H K

1561 AGCTACCCGCTCGAAATATTATTGAGTCACAGTGGATAACATTACAATTACTTTATTG 1620
-----+-----+-----+-----+-----+-----+-----+-----+
TCGATGGCGGAGCTTTATAATAACTCAGTGTACCTATTGTAAATGTTAATTGAAATAAAC
a S Y P L E I L L S H S G - H Y N - L Y L
b A T R S K Y Y - V T V D N I T I N F I C
c L P A R N I I E S Q W I T L Q L T L F A

Figure 11I

1621

ந டு உ

1691

ਸੁ ਧੁ ਪ

1740

ந டு ப

Figure 11J

AGCATGTAACAATTACAAGACGATGAAATCAATTCGTTTCATGGTCGATCAGCAG
-----+-----+-----+-----+-----+ 1860
TCGTACATTTGTAAATGTTCTGCTACTTTTAGTTAAAGCAAGTAACCAGCTAGTGCCTC

1801

a S M - N I Y K T M K I N F V H W S I T Q
b A C K T F T R R - K S I S F I G R S R S
c H V K H L Q D D E N Q F R S L V D H A A

CGCATCATATAAATAAGTATAGAACAGGTTCTGTTGGCTGGTTCGACCAACAGTCGG
-----+-----+-----+-----+-----+ 1920
GCGTAGTATAATTTTATCATATCTTGTTCCAAGACAACCGGACCAAGCTGGTGTGCAGCC

1861

a R I I L K I V - N R F C W P G S T N S R
b A S Y - K - Y R T G S V G L V R P T V G
c H H I K N S I E Q V L L A W F D Q Q S V

TAGACAGTGTATGTGCCATCGTCTGGCACGCCAGGCCAGGCTATGGCGGAAGAGGAG
-----+-----+-----+-----+-----+ 1980
ATCTGTCAACAATACACGGTAGCAGACCGTGCGGTCCGGTCCGATACCGCCTTCTCCCTC

1921

a - T V L C A I V W H A R P R L W R K R E
b R Q C Y V P S S G T P G H G Y G G R G S
c D S V M C H R L A R Q A T A M A E E G A

Figure 11K

1981	2040
CGCTTTATTGCGTATTCA	CGCTTTATTGCGTATTCA
TCCTGAAAGAGGCAT	TCCTGAAAGAGGCAT
TGATGCGAGAACTTT	TGATGCGAGAACTTT
TTTGGCAAGC	TTTGGCAAGC
GC	GC
CGGAAATAAACGCATAAGTAGGACTTTT	CGGAAATAAACGCATAAGTAGGACTTTT
TCTCCGTAAC	TCTCCGTAAC
TACGCTCTTTGAAACCGTT	TACGCTCTTTGAAACCGTT
TCG	TCG

a R F I C V F I L K K R H - C E K L L A S
b A L F A Y S S - K R G I D A R N F W Q A
c L Y L R I H P E K E A L M R E T F G K R

2040
GGTTACGTTGATTATCGAGCCTGTTTCTCTCCGATCAGGCTGAACCTTCTCTCAACAC
-----+-----+-----+-----+-----+-----+ 2100
CCAAATGCAACTAATAGCTCGGACCAAAGAGAGGGCTAGTCCGACTTGAAAGGAGTTGTG

a G L R - L S S L V S L P I R L N F P Q H
b V Y V D Y R A W F L S R S G - T F L N T
c F T L I I E P G F S P D Q A E L S S T R

2101
GATATGCCGTTGAATTTTCACTTTCCTCGTCATTTCAACGCTTACTGAAATGGTTACGTA
-----+-----+-----+-----+-----+-----+ 2160
CTATACGGCAACTTAAAAGTGAAGAGCAGTAAAGTTGCGCAATGACTTTACCAATGCAT

a D M P L N F H F L V I S T R Y - N G Y V
b I C R - I F T F S S F Q R V T E M V T -
c Y A V E F S L S R H F N A L L K W L R N

Figure 11L

2161

ਥੁ ਧੁ ਧੁ

2221

ந டு உ

2281

ത മ ഉ

Figure 11M

2341

a I G M A L S G K L R E N Y K F T D A E K
b - E W R - V V N C A K I I N S L M L R N
c R N G A E W - T A R K L - I H - C - E T

2401

a L E R R Q Q A L L R L I K Q I Q E D N G
b W S A D S R L C C V - N K Y R R I M G
c G A Q T A G F A A F D K T N T G G - W G

2461

a A T L R P L T E E N S D P D L Q N A Y Q
b Q R C V R L P K R I V I L I Y R M R I K
c N V A S A Y R R E - - S - F T E C V S N

Figure 11N

2521
ATTATCGCTCTTGCAATGGCGCTTACTGCCGCGGTTGTCAAAAAGAAAAACGGAT
-----+-----+-----+-----+-----+-----+ 2580
TAATAGCGAGAACGTTACCGCGAATGACGGCCGCCAAGTTTTCCTTTTTCGGCTA

a I I A L A M A L T A G G L S K K K R D
b L S L L Q W R L L P A G C Q K R K N A I
c Y R S C N G A Y C R R V V K K E K T R F

2581
TTGCAATCGCAACTGGATACGTTACAGCGGAGGATGGAACTTGCCGTTTTAGTT
-----+-----+-----+-----+-----+-----+ 2640
AACGTTAGCGTTGACCTATGCAATGTCGCCTCCTCCCTACCCCTTGAACGGCAAAATCAA

a L Q S Q L D T L Q R R R D G N L P F L V
b C N R N W I R Y S G G M G T C R F - F
c A I A T G Y V T A E E G W E L A V F S L

2641 TACTGGAACTTGGCGAAGTGGATACCGTACGCTGTCTCTGAAGCGTTTATGCAACA
 -----+-----+-----+-----+-----+-----+ 2700
 ATGACCTTGAACCGCTTCACCTATGGCATGCGACAGGAGACTTCGCAAAATACGTTGT

a Y W N L A K W I P Y A V L S E A F Y A T
b T G T W R S G Y R T L S S L K R F M Q Q R
c L E L G E V D T V R C P L - S V L C N R

Figure 110

2881 GTTATTCCTTGGCCTTGAAAAAGAGTGCCAGCGTGAGGAGTGGAATTGCCAGTTGCCGCC
-----+-----+-----+-----+-----+-----+-----+ 2940
CAATAAGGAACCGGAACCTTTTCTCAGGTCGCACTCCTCACCTAAACGGTCAACGGCGG

a V I P W P - K R V P A - G V D L P V A A
b L F L G L E K E C Q R E E W I C Q L P P
c Y S L A L K K S A S V R S G F A S C R L

2941 TAATACATTACTGCCGCTACTACTCGATATTATTGTGAGCGCTGGCTTTTCAGTGATTG
-----+-----+-----+-----+-----+-----+ 3000
ATTATGTAATGACGGCGATGATGAGCTATAATAAACACTCGCGACCGAAAGTCACTAAC

a - Y I T A A T T R Y Y L - A L A F Q - L
b N T L L P L L L D I I C E R W L F S D W
c I H Y C R Y Y S I L F V S A G F S V I G

3001 GTTGCTTGATAGACTTACCGCTATAGTTTCTTCATCGAAGATGTTCAATCGGTTACTCCA
-----+-----+-----+-----+-----+-----+ 3060
CAACGAACTATCTGAATGCGGATATCAAGAAGTAGCTTCTACAAGTTAGCCAATGAGGT

a V A - - T Y R Y S F F I E D V Q S V T P
b L L D R L T A I V S S S K M F N R L L Q
c C L I D L P L - F L H R R C S I G Y S N

Figure 11Q

```

2701      GCGATAGACAACGATGAAATGCCCTTATCGCAGTGGTTCAGACGCGTGCGAGACTGGCC      2760
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCGCTATCTGTTGCTACTTTACGGGAATAGCGTCACCAAGTCTGCGCACCCGTCTGACCCGG

a      G D R Q R - N A L I A V V Q T R G R L A
b      A I D N D E M P L S Q W F R R V A D W P
c      R - T T M K C P Y R S G S D A W Q T G R

2761      GGATCGCTGTGAACGGTCCGTATTTTGCTAAGAGCAGTAGCCTTTGAACCTTAGCATATG      2820
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCTAGCGACACTTGCCCGGCATATAAACGATTCTCGTCATCGGAACCTTGAATCGTATAC

a      G S L - T G P Y F A K S S S L - T - H M
b      D R C E R V R I L L R A V A F E L S I C
c      I A V N G S V F C - E Q - P L N L A Y A

2821      CATCGAACCCCTCGGAGCAAAGTCGTTTGGCCCGCAGCATTAGTACGTTTGCCTTGCT      2880
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTAGCTTGGAGCCCTCGTTTCAGCAAACCGGCGTCGTAATCATGCAAAACGCGACGCAACGA

a      H R T L G A K S F G R S I S T F A S F A
b      I E P S E Q S R L A A A L V R L R R L L
c      S N P R S K V V W P Q H - Y V C V V C C

```

Figure 11P

3061 ACAAATTGATGCGCAGTTTATGCTGATACCCGATAACTGTTTAAACGACGAAGATCAACG 3120
-----+-----+-----+-----+-----+-----+-----+
TGTTGAACTACGCGTCAAATACGACTATGGGCTATTGACAAAATTGCTGCTTCTAGTTGC
a T T - C A V Y A D T R - L F - R R R S T
b Q L D A Q F M L I P D N C F N D E D Q R
c N L M R S L C - Y P I T V L T T K I N V
3121 TGAACAAATTCTCGAAACGCTTCGTGAAGTAAAGATAAATCAGGTTTATTCTGATACCT 3180
-----+-----+-----+-----+-----+-----+-----+
ACTTGTTAAGAGCTTTGCCGAAGCACTTCTATTCTATTAGTCCAAAATAAGACTATGGA
a - T N S R N A S - S K D K S G F I L I P
b E Q I L E T L R E V K I N Q V L F - Y L
c N K F S K R F V K - R - I R F Y S D T W
3181 GGCTTCAATATTAGGTAAATTGGCTTTCTGGCTCATCATGAGGCGTCAGGATGGATTG 3240
-----+-----+-----+-----+-----+-----+-----+
CCGAAAGTTATAAATCCATTTAACCGAAAGACCGAGTAGTACTCCGCAGTCCCTACCTAAC
a G F Q Y L G K L A F W L I M R R Q D G L
b A F N I - V N W L S G S S - G V R M D W
c L S I F R - I G F L A H H E A S G W I G

Figure 11R

```

3241      GGATCTCATTACTGAACGTAATATTCAGCTTTTATTCAATTAGCAGATTAGCTGAACG      3300
      -----+-----+-----+-----+-----+-----+-----+-----+
      CCTAGAGTAATGACTTGCAATTATAAGTCGAAAAAATAAGTTAATCGTCCTAATCGACTTGC

      a      G S H Y - T - Y S A F Y S I S R I S - T
      b      D L I T E R N I Q L F I Q L A G L A E R
      c      I S L L N V I F S F L F N - Q D - L N G

3301      GCCTTTAGCAACCAATATGTTCTGGCGCAAGGACAATATGAAACTATCATAACGGTCGT      3360
      -----+-----+-----+-----+-----+-----+-----+-----+
      CGGAAATCGTTGGTTATACAAGACCGCGTTCCTGTTATACCTTTGATAGTATTGCCAGCA

      a      A F S N Q Y V L A A R T I - N Y H N G R
      b      P L A T N M F W R Q G Q Y E T I I T V V
      c      L - Q P I C S G G K D N M K L S - R S Y

3361      ATTCTCTTATGTCAGATACTCAAGCAACCTTCTTAGACGAAGAACTGCTTTTAAAGCG      3420
      -----+-----+-----+-----+-----+-----+-----+-----+
      TAAGAGAAATACAGTCTATGAGTTCGTTTGGAAGAATCTGCTTCTTGACGAAAAAATTTCCG

      a      I L L C Q I L K Q T F L D E E L L F K A
      b      F S Y V R Y S S K P S - T K N C F L K R
      c      S L M S D T Q A N L L R R R T A F - S V

```

Figure 11S

3421 TTGGCTAACTGGAACCCGACGCTTCCAGGTATTCCCTCAACGATTATTTTGTTCGC
-----+-----+-----+-----+-----+-----+ 3480
AACCGATTGACCTTTGGCGTCGCAAGGTCCCATAGGAGTTGCTAATAAAACAACGCG

a L A N W K P A A F Q G I P Q R L F L L R
b W L T G N P Q R S R V F L N D Y F C C A
c G - L E T R S V P G Y S S T I I F V A R

3481 GATGGCTTGCAATGAGTTGTTCTCCACCTCTTCCAGCTCCGCCGAGCTCTGGTTACGA
-----+-----+-----+-----+-----+-----+ 3540
CTACCCGAACGTTACTCAACAAGAGGTGGAGAAAGGTCGAGCGGCTCGAGACCAATGCT

a D G L A M S C S P P L S S S A E L W L R
b M G L Q - V V L H L F P A P S S G Y D
c W A C N E L F S T S F Q L R R A L V T I

3541 TTACATCATCGACAAATAAAATTTTCNTGGAGTCGCAATGCGTTTCATGTTAGTGAGGGA
-----+-----+-----+-----+-----+-----+ 3600
AATGTAGTACGTGTTTATTTTAAAGNACCTCAGCGTTACGCAAGTACCAATCCACTCCCT

a L H H R Q I K F X G V A M R S W L G E G
b Y I I D K - N F X E S Q C V H G - V R E
c T S S T N K I X W S R N A F M V R - G S

Figure 11T

GTCAGGGCGCAACAGTGGCTCAGTGTATGCGCGGTCGGCAGGATATGGTTCTGGCGACG
-----+-----+-----+-----+-----+-----+ 3660
CAGTCCCGCGTTGTACCGAGTCACATACGCGGCCAGCCGTCCTATACCAAGACCGCTGC

3601

start lcrD*
a V R A Q Q W L S V C A G R Q D M V L A T
b S G R N S G S V Y A R V G R I W F W R R
c Q G A T V A Q C M R G S A G Y G S G D G

GTGTTAATAATCGCTATTGTGATGATGCTGTACCCCTTGCCGACCTGGATGGTTGATATC
-----+-----+-----+-----+-----+ 3720
CACATAATTAGCGATAACACTACTACGACAATGGGAACGGCTGGACCTACCAACTATAG

3661

a V L L I A I V M M L L P L P T W M V D I
b C Y - S L L - - C C Y P C R P G W L I S
c V I N R Y C D D A V T L A D L D G - Y P

CTGATTACTACAACCTTATGTTTTCAGTGATCCCTGCTCTTAATTGCTATTATCTTAGT
-----+-----+-----+-----+-----+ 3780
GACTAATGATAGTTGGAATACAAAGTCAGGACGAGAAATTAACGATAAATAGAATCA

3721

a L I T I N L M F S V I L L L I A I Y L S
b - L L S T L C F Q - S C S - L L F I L V
c D Y Y Q P Y V F S D P A L N C Y L S - -

Figure 11U

```

3781      GACCCCTCGATTATCGGTATTCCGCTTTACTTATTACTATTACTATATATCGTTTG      3840
-----+-----+-----+-----+-----+-----+-----+-----+
CTGGAGAGCTAAATAGCCATAAAGGCAGAAATAATGAATAATGATGTAATATAGCAAAC

      a   D P L D L S V F P S L L L I T T L Y R L
      b   T L S I Y R Y F R L Y Y L L L H Y I V C
      c   P S R F I G I S V F I T Y Y I I S F V

3841      TCACTCACAATCAGCACATCACGGCTGGTACTGTACACATAATGCCGGTAATATTGTG      3900
-----+-----+-----+-----+-----+-----+-----+
AGTGAGTGTAGTCGTAGTGCCGACCATGACAATGTTGTATTACGGCCATTATAACAC

      a   S L T I S T S R L V L L Q H N A G N I V
      b   H S Q S A H H G W Y C Y N I M P V I L W
      c   T H N Q H I T A G T V T T - C R - Y C G

3901      GATGCTTTCGGTAAGTTTGTCTAGGAGGAAATCTCACCGTTGGTTCGTATTACC      3960
-----+-----+-----+-----+-----+-----+-----+
CTACGAAAGCCATTCAAACAGCATCCTCTTTAGAGTGGCAACCAACCAAGCATAAATGG

      a   D A F G K F V V G G N L T V G L V V F T
      b   M L S V S L S - E E I S P L G W S Y L P
      c   C F R - V C R R R K S H R W V G R I Y H

```

Figure 11V

a I I T I V Q F I V I T K G I E R V A E V
b S L L S C N L L S L Q K V S R G W R K L
c H Y Y R A I Y C H Y K R Y R E G G G S -

a S A R F S L D G M P G K Q M S I D G D L
b A H V S R L M G C Q A N K - V S M A I C
c R T F L A - W D A R Q T N E Y R W R F A

a R A G V I D A D H A R T L R Q H V Q Q E
b V P E L S M Q T M P V H - D S M S S R K
c C R S Y R C R P C P Y I K T A C P A G K

Figure 11W

AGCCGCTTCTCGGTGCGATGGACGGTGCATGAAATTTGTTAAAGCGGATACGATGCCC
-----+-----+-----+-----+-----+-----+ 4200
TCGGCGAAGAGCCACGCTACCTGCCACGCTACTTTAAACAATTTCCGCTATGCTAACGG

4141

a S R F L G A M D G A M K F V K G D T I A
b A A F S V R W T V R - N L L K A I R L P
c P L S R C D G R C D E I C - R R Y D C R

GGTATTATTGTTCTGTTGAACATTATCGCGGTATCATTATCGCTATCGTACAATAT
-----+-----+-----+-----+-----+ 4260
CCATAATAACAAGACCACCTTGTAATAGCCGCATAGTAATAGCGATAGCATGTTATA

4201

a G I I V V L V N I I G G I I I A I V Q Y
b V L L L F W - T L S A V S L S L S Y N M
c Y Y C C S G E H Y R R Y H Y R Y R T I -

GATATGTCGATGAGTGAGGCTGTTCACTTATAGCGTACTGTCAATCGGAGATGGTTTA
-----+-----+-----+-----+-----+ 4320
CTATACAGCTACTCACTCCGACAAGTGTAATATCGCATGACAGTTAGCCTCTACCAAAT

4261

a D M S M S E A V H T Y S V L S I G D G L
b I C R - V R L F T L I A Y C Q S E M V Y
c Y V D E - G C S H L - R T V N R R W F M

Figure 11X

TGTGGGCAAATTCCATCGCTGCTGATTCCCTTAGCGGGGAATTATTGTACCCGTGTC
-----+-----+-----+-----+-----+ 4380
ACACCCGTTTAAGGTAGCAGCAGCTAAAGGGAATCGCGCCCTTAATAACAGTGGGCACAG

4321

a C G Q I P S L L I S L S A G I I V T R V
b V G K F H R C - F P L A R E L L S P V S
c W A N S I A A D F P - R G N Y C H P C P

CCGGGTGAGAAACGCCAGAACCTGGCGACAGAGTTGAGTTCTCAAATTGCCAGACAACCT
-----+-----+-----+-----+-----+ 4440
GGCCCACTCTTTGGGTCTTGGACCGCTGCTCAACTCAAGAGTTTAACGGTCTGTTGGA

4481

a P G E K R Q N L A T E L S S Q I A R Q P
b R V R N A R T W R Q S - V L K L P D N L
c G - E T P E P G D R V E F S N C Q T T S

CAGTCGCTCATATTAACCGCTGTGGTTTAAATGCTCCTCGCTTAAATTCCTGGCTTTCCT
-----+-----+-----+-----+-----+ 4500
GTCAGCGAGTATAATTGGCGACACCAAAATTACGAGGAGCGAAATTAAAGGACCGAAAGGA

4441

a Q S L I L T A V V L M L L A L I P G F P
b S R S Y - P L W F - C S S L - F L A F L
c V A H I N R C G F F N A P R F N S W L S F

Figure 11Y

TTTATCACTCGCTTTCTTTTTCAGCGTTGTAGCATGCGCAATTATCCTCATTCGCCGC
-----+-----+-----+-----+-----+-----+ 4560
AAATAGTGAGAGCGAAAGAAAAGTCGCAACAATCGTAACGGTTAATAGGAGTAAGCGGCG

4501

a F I T L A F F S A L L A L P I I L I R R
b L S L S L S F Q R C - H C Q L S S F A A
c Y H S R F L F S V V S I A N Y P H S P Q

Tn insertion P11C3
⇕

AAAAAGTCTGTGTTTCCGCAAAATGGCGTCGAAGCACCGGAAAAAGATAGTATGTTCCC
-----+-----+-----+-----+-----+-----+ 4620
TTTTTCAGACACCAAGCGTTTACCGCAGCTTCGTGGCCTTTTCTATCATACCAAGGG

4561

a K K S V V S A N G V E A P E K D S M V P
b K S L W F P Q M A S K H R K K I V W F P
c K V C G F R K W R R S T G K R - Y G S R

GGCGCATGTCCTCTAATCTTACGTCTTAGCCCGACGTTACATTCTGCCGACCTGATTCTGT
-----+-----+-----+-----+-----+-----+ 4680
CCGCGTACAGGAGATTAGAATGCAGAATCGGGCTGCAATGTAAGACGGCTGGACTAAGCA

4621

a G A C P L I L R L S P T L H S A D L I R
b A H V L - S Y V L A R R Y I L P T - F V
c R M S S N L T S - P D V T F C R P D S -

Figure 11Z

4681 GATATTGACGCCATGAGATGGTTTTTTATTGAGGATACGGCGTCCCTCTCCCTGAGGTG 4740
 -----+-----+-----+-----+-----+-----+-----+
 CTATAACTGCGGTACTCTACCAAAAATAAACTCCTATGGCCGAGGAGAGGGACTCCAC

 a D I D A M R W F L F E D T G V P L P E V
 b I L T P - D G F Y L R I P A S L S L R -
 c Y - R H E M V F I - G Y R R P S P - G E

 4741 AATATTGAGGTTTGCTGAACCCACCGAAAAATTGACGGTACTGCTATATCAGGAACCC 4800
 -----+-----+-----+-----+-----+-----+-----+
 TTATAACTCCAAAACGGAAGTGGGTGGCTTTTAACTGCCATGACGATATAGTCTTGGG

 a N I E V L P E P T E K L T V L L Y Q E P
 b I L R F C L N P P K N - R Y C Y I R N P
 c Y - G F A - T H R K I D G T A I S G T R

 4801 GTATTAGTTTATCTATTCCTCAGGCGGATTATTATTGATAGGCGGACGCTAGT 4860
 -----+-----+-----+-----+-----+-----+-----+
 CATAAATCAATAGATAAGGCGAGTCCGCCTAATAAATACTATCCGCGCTGCGATCA

 a V F S L S I P A Q A D Y L L I G A D A S
 b Y L V Y L F P L R R I I Y - - A R T L V
 c I - F I Y S R S G G L F I D R R G R - C

Figure 11AA

4920

U

4980

U

5040

C

Figure 11AB

5041

൯൧൦

5101

ਭ ਯੁ ਪ

5161

ਭ ਭ ਭ

Figure 11AC

Figure 11AD

5401 AAGACGCAGATCCTGCAACTTATCGAGCAGCGCTGAAGCAGTCAGCCAAATTATTCATT 5460
-----+-----+-----+-----+-----+-----+-----+
TTCTGCGTCTAGGACGTTGAATAGCTCGTCCGCGACTTCGTCAGTCGGTTTAATAAGTAA

a K T Q I L Q L I E Q A L K Q S A K L F I
b R R R S C N L S S R R - S S Q P N Y S L
c D A D P A T Y R A G A E A V S Q I I H C

5461 GTCAC TTCTGTCACACCCGACGTTTCTTTCGCGAAAATTACAGAAGCCACCTTGTTCGAC 5520
-----+-----+-----+-----+-----+-----+-----+
CAGTGAAGACAGCTGTGGCTGCAAGAACGCTTTTAAATGTCTTCGGTGGAAACAAGCTG

a V T S V D T R R F L R K I T E A T L F D
b S L L S T P D V S C E K L Q K P P C S T
c H F C R H P T F L A K N Y R S H L V R R

5521 GTACCGATTTTGTTCATGGCAGGAATTAGGAGAGGAGAGCCCTTATACAAGTGGTAGAAAGT 5580
-----+-----+-----+-----+-----+-----+-----+
CATGGCTAAAACAGTACCGTCCTTAATCCTCTCCTCTCGGAATATGTTCAACCATCTTTCA

a V P I L S W Q E L G E E S L I Q V V E S
b Y R F C H G R N - E R R A L Y K W - K V
c T D F V M A G I R R G E P Y T S G R K Y

Figure 11AE

ATTGACCTTAGCGAAGAGGAGTTGGCGACAATGAAGAATGAATTGATGCAACGTCTGAG 5640

5581

-----+-----+-----+-----+-----+-----+-----+
TAACTGGAATCGCTTCTCCTCAACCGCCTGTTACTTCTTACTTAACTACGTTCAGACTC

	end lcrD*										start yscN*?									
a	I	D	L	S	E	E	L	A	D	N	E	E	-	I	D	A	T	S	E	
b	L	T	L	A	K	R	S	W	R	T	M	K	N	E	L	M	Q	R	L	R
c	-	P	-	R	R	G	V	G	G	Q	-	R	M	N	-	C	N	V	-	G

GCTGAAATATCCGCCCCCGATGGTTATTGTCGATGGGGCCGAATTCAGGATGTCAGCGC 5700

5641

-----+-----+-----+-----+-----+-----+-----+
CGACTTTATAGCGGGGGCTACCAATAACAGCTACCCCGGCTTAAGTCCCTACAGTCGCG

a	A	E	I	S	A	P	R	W	L	L	S	M	G	P	N	S	G	C	Q	R
b	<u>L</u>	K	Y	P	P	P	D	G	Y	C	R	W	G	R	I	Q	D	V	S	A
c	-	N	I	R	P	P	M	V	I	V	D	G	A	E	F	R	M	S	A	Q

AACGTTGTTAAATCGTGTTGCCCTGGGGTATTATGGGCGAGTTGTGCTGTATAAGCC 5760

5701

-----+-----+-----+-----+-----+-----+-----+
TTGCAACAATTACGCACCAACGGACCCCATATAATACCCGCTCAACACGACATATTTCGG

a	N	V	V	K	C	V	V	A	W	G	I	Y	G	R	V	V	L	Y	K	A
b	T	L	L	N	A	W	L	P	G	V	F	M	G	E	L	C	C	I	K	P
c	R	C	-	M	R	G	C	L	G	Y	L	W	A	S	C	A	V	-	S	L

Figure 11AF

5761 TGGAGAGAACTTGCTGAAGTCGTGGGATTAATGGCAGCAAGCTTTGCTATCTCCTTT 5820
-----+-----+-----+-----+-----+-----+
ACCTCTTCTTGAACGACTTCAGCACCCCTAATTACCGTCGTTTCGAAACGATAGAGGAAA

YSCN*
a W R R T C - S R G D - W O Q S F A I S F
b G E E L A E V V G I N G S K A L L S P F
c E K N L L K S W G L M A A K L C Y L L L
5821 TACGAGTACAATCGGGCTTCACTCGGGCAGCAAGTGTGATGGCCCTTAAGCGACGCCATCAG 5880
-----+-----+-----+-----+-----+-----+
ATGCTCATGTTAGCCCGAAGTGACGCCCGTCGTTCACTACCGGAATTGCGTGGGTAGTC

a Y E Y N R A S L R A A S D G L K R R H Q
b T S T I G L H C G Q Q V M A L S D A I R
c R V Q S G F T A G S K - W P - A T P S G
5881 GTTCCCGTGGCGAAGCGTTATTAGGGCGAGTTATTGATGGCTTTGGTCGCCCTTGAT 5940
-----+-----+-----+-----+-----+-----+
CAAGGCACCCGCTTCGCAATAATCCCGCTCAATACTACCGAAACCAGCAGGGGAACCTA

a V P V G E A L L G R V I D G F G R P L D
b F P W A K R Y - G E L L M A L V V P L M
c S R G R S V I R A S Y - W L W S S P - W

Figure 11AG

GGCCGCGAACTGCCGACGCTGCTGTGGAAGACTATGATGCAATGCCCTCCTCCGCAATG 6000
-----+-----+-----+-----+-----+-----+
CCGGCGCTTGACGGCTGCAGACGACCTTTCTGATACTACGTTACGGAGGAGGCGGTAC

5941

a G R E L P D V C W K D Y D A M P P P A M
b A A N C P T S A G K T M M Q C L L P Q W
c P R T A R R L L E R L - C N A S S R N G

GTTGACAGCCATCACTCAACCAATTAATGACGGGATTCGCGCTATTGATAGCGTTGCG 6060
-----+-----+-----+-----+-----+-----+
CAAGCTGTCGGATAGTGAGTTGGTAATTACTGCCCCCTAAGCGCGATACTATCGCAACGC

6001

a V R Q P I T Q P L M T G I R A I D S V A
b F D S L S L N H - - R G F A L L I A L R
c S T A Y H S T I N D G D S R Y - - R C D

ACCTGTGCGAAGGCAACGAGTGGGTAATTTTCTGCTCCTGGCGTGGGAAAAGCACG 6120
-----+-----+-----+-----+-----+-----+
TGGACACCGCTTCCCGTTGCTCACCCATAAAAAGACGAGGACCGCACCCCTTTTCGTGC

6061

a T C G E G Q R V G I F S A P G V G K S T
b P V A K G N E W V F F L L L A W G K A R
c L W R R A T S S G Y F F C S W R G E K H A

Figure 11AH

CTTCTGGCGATGCTGTGTAATGCGCCAGACGACAGCAATGTTCTGGTGAATTGGT 6180
-----+-----+-----+-----+-----+-----+
GAAGACCGCTACGACACATTACGCGGTCTGCGTCTGCTGTTACAAGACCACAATTAACCA

6121

a L L A M L C N A P D A D S N V L V L I G
b F W R C C V M R Q T Q T A M F W C - L V
c S G D A V - C A R R R Q Q C S G V N W -

GAACGTGACGAGAAGTCCGCGAATTATCGATTTTACACTGTCTGAAGAGACCCGAAAA 6240
-----+-----+-----+-----+-----+-----+
CTTGACCTGCTCTTCAGGCGCTTAAGTAGCTAAATGTGACAGACTTCTCTGGGCTTTT

6181

a E R G R E V R E F I D F T L S E E T R K
b N V D E K S A N S S I L H C L K R P E N
c T W T R S P R I H R F Y T V - R D P K T

CGTTGTGTCATTGTTGCGCAACCTCTGACAGACCCGCCCTTAGAGCGGTGAGGGCGCTG 6300
-----+-----+-----+-----+-----+-----+
GCAACACAGTAACAACAGCGTTGGAGACTGTCTGGCGGAATCTCGCGCACTCCCGCGAC

6241

a R C V I V V A T S D R P A L E R V R A L
b V V S L L S Q P L T D P P - S A - G R C
c L C H C C R N L - Q T R L R A R E G A V

Figure 11AI

6301

π

د

U

6361

α

د

U

6421

၈၄

2

U

Figure 11AJ

ACGGGAATGGGAGAAAAGGCAGTATTACCGCATTTTATACGGTACTGGTGGAAGCGGAT 6540
-----+-----+-----+-----+-----+-----+-----+
TGCCCTTACCCCTCTTTTCCGTCATAATGGCGTAAATAATGCCATGACCACCTTCCGCTA

a T G M G E K G S I T A F Y T V L V E G D
b R E W E K K A V L P H F I R Y W W K A M
c G N G R K R Q Y Y R I L Y G T G G R R -

GATATGAATGAGCCGTTGGCGGATGAAGTCCGTTCACTGCTTGATGGACATATTGTACT 6600
-----+-----+-----+-----+-----+-----+-----+
CTATACTTACTTCGGCAACCGCCTACTTTCAGGCAAGTGACGAACCTACCTGTATAACATGA

YSCN*

a D M N E A V G G - S P F T A - W T Y C T
b I - M K P L A D E V R S L L D G H I V L
c Y E - S R W R M K S V H C L M D I L Y Y

ATCCCGACGGCTTGCAGAGAGGGGCATTATCCGCCATTGACGTGTGGCAACGCTCAG 6660
-----+-----+-----+-----+-----+-----+-----+
TAGGGCTGCCGAACGTCTCTCCCCCGTAATAGGACGGTAAGTGCACAACCGTTGCCGAGTC

a I P T A C R E G A L S C H - R V G N A Q
b S R R L A E R G H Y P A I D V L A T L S
c P D G L Q R G G I I L P L T C W Q R S A

Figure 11AK

6661 CCGCGTTTCCAGTCGTTACAGCCATGAGCATCGTCAACTGGCGGCGATATTGGACG 6720
-----+-----+-----+-----+-----+-----+
GGCGCAAAAGGTCAGCAATGGTCGTTACTCGTAGCAGTTGACCGCCGCTATAACGCTGC

a P R F S S R Y Q P - A S S T G G D I A T
b R V F P V V T S H E H R Q L A A I L R R
c A F F Q S L P A M S I V N W R R Y C D G

6721 GTGCCGTGGCGCTTACAGGAGGTTGAACGTGTTAATACGCATTGGGGAATACCAGCGAGG 6780
-----+-----+-----+-----+-----+-----+
CACGGACCGCGAAATGGTCCTCCAACTTGACAATTATGCGTAACCCCTTATGTCGCTCC

a V P G A L P G G - T V N T H W G I P A R
b C L A L Y Q E V E L L I R I G E Y Q R G
c A W R F T R R L N C - Y A L G N T S E E

6781 AGTGATACAGATACTGACAAAGCCATTGATACCTATCCGGATATTGCCACATTTTGGC 6840
-----+-----+-----+-----+-----+-----+
TCAACTATGTCTAGACTGTTTCGGTAACCTATGATAGGCCCTATAAACGTGTAAAAACGC

a S - Y R Y - Q S H - Y L S G Y L H I F A
b V D T D T D K A I D T Y P D I C T F L R
c L I Q I L T K P L I P I R I F A H F C D

Figure 11AL

ACAAAGTAAGGATGAAGTATGCGGACCCGAGCTACTTATAGAAAAATTACACCAATACT
-----+-----+-----+-----+-----+-----+ 6900
TGTTTCATTCCTACTTCATACGCCCTGGCTCGATGAATATCTTTTAAATGTGGTTTATGA

6841

end yscN*
a T K - G - S M R T R A T Y R K I T P N T
b Q S K D E V C G P E L L I E K L H Q I L L
c K V R M K Y A D P S Y L - K N Y T K Y S

CACCGAGTGATCATGGAACCTTGTCTGGAGATAATCGCGGCTGAAAAGCAATTACGCG
-----+-----+-----+-----+-----+ 6960
GTGGCTACTAGTACCTTTGAAACGACCTCTATTAGCGCGCGGACTTTTCGTTAATGCGC

6901

yscO*

a H R V I M E T L L E I I A R L K S N Y A
b T E - S W K L C W R - S R G - K A I T R
c P S D H G N F A G D N R A A E K Q L R G

GCAAGCTTACCGTACTTGATCAGCAGCAACAGCGGATTATTACGGAACAGCAGATTGCC
-----+-----+-----+-----+-----+ 7020
CGTTCGAATGGCATGAACCTAGTCGTGTTGTCCGCTAATAATGCCCTTGTCTCTAAACGG

6961

a A S L P Y L I S S N R R L L R N S R F A
b Q A Y R T - S A A T G D Y Y G T A D L P
c K L T V L D Q Q Q Q A I I T E Q Q I C Q

Figure 11AM

7080

U

7140

U

7200

U

Figure 11AN

CGAAGCGAATTACAGAAGAATTTTAATGCGCTTATGAAAAGAAAATAACTATG 7260
-----+-----+-----+-----+-----+-----+
GCTTCGCTTAATGTCTTCTTAAATACGCGAATACTTTTCTTTCTTTTAAATGATAC

7201

end yscO*

a R S E L Q K N F N A L M K K K E K I T M
b E A N Y R R I L M R L - K R K K L L W
c K R I T E E F - C A Y E K E R K N Y Y G

GTATTAAGCGATGCGTATTACCAAAGTTGAGGGAAGTCTTGGGTTGCCAGTCTTA 7320
-----+-----+-----+-----+-----+-----+
CATAATTGCTACGCATAATGGTTTCAACTCCCTTCAGAACCCAACGGTACGGTCAGAAT

7261

start yscP*

a V L S D A Y Y Q S - G K S W V A M P V L
b Y - A M R I T K V E G S L G L P C Q S Y
c I K R C V L P K L R E V L G C H A S L I

TCAGGATGATAACGAGGCGGAGGCGGAACGTATGGACTTTGAACAACATCATGCACGCG 7380
-----+-----+-----+-----+-----+-----+
AGTCCTACTATTGCTCCGCCCTCCGCCCTTGACATACCTGAAACTTGTGAGTACGTGTC

7321

a S G - - R G G G T Y G L - T T H A P G
b Q D D N E A E A E R M D F E Q L M H Q A
c R M I T R R R R N V W T L N N S C T R H

Figure 11AO

7381

a I T H W - E - S S C S I E - E R G F H A
b L P I G E N N P P A A L N K N V V F T Q
c Y P L V R I I L L Q H - I R T W F S R N

7441

a T L S C - W R L S - R C R V - S M - I R
b R Y R V S G Y L D G V E C E V C E S G
c V I V L V A V I L T V - S V K Y V N Q G

7501

a G A N P V K N Q C P S S - N L P F D E S
b G L I Q L R I N V P H E I Y R S M K A
c G - S S - E S M S L I M K F T V R - K R

Figure 11AP

7561 GCTAAAGCAGTGGCTGGAGTCTCAGTTGCTGCATATGGGGTATATAATTTCCTGGAGAT 7620
-----+-----+-----+-----+-----+-----+-----+
CGATTTCGTCACCGACCTCAGAGTCAACGACGTATACCCCATATATAAAGGACCTCTA

a A K A V A G V S V A A Y G V Y N F P G D
b L K Q W L E S Q L L H M G Y I I S L E I
c - S S G W S L S C C I W G I - F P W R Y

7621 ATTCTATGTTAAGAATAGCGAATGAAGAGCGTCCGTGGTGAGATACTTCCAACGCAAG 7680
-----+-----+-----+-----+-----+-----+-----+
TAAGATACAATTCTTATCGCTTACTTCTCGCAGGCACCCACCTCTATGAAGTTGCGTTC

end yscP* start yscQ*?

a I L C - E - R M K S V R G W R Y F Q R K
b F Y V K N S E - R A S V G G D T S N A R
c S M L R I A N E E R P W V E I L P T Q G

7681 GCGCTACCATTTGGTGAGCTGACATTGAGTATGCAACAATATCCAGTACAGCAAGGACAT 7740
-----+-----+-----+-----+-----+-----+-----+
CGCGATGGTAACCACTCGACTGTAACTCATACGTTGTTATAGGTCATGTCGTTCCCTGTA

start yscQ*?

a A L P L V S - H - V C N N I Q Y S K G H
b R Y H W - A D I E Y A T I S S T A R D I
c A T I G E L T L S M O O Y P V Q Q G T L

Figure 11AQ

7741

ਥੁ ਧੁ ਧੁ

7801

ந டு ப

7861

ਥੁ ਮੁ ਪ

Figure 11AR

GTCAGAACGAGCCGAACATCCTGCAGTAATCTACCATCAGCTAGCGTGCATATTA
-----+-----+-----+-----+-----+-----+ 7980
CAGTCTTGCTCGCGGTTGTAGGACGTCATTAGATGGTGTAGTCGATCGCAACGTATAAT

7921

a V R T S R Q H P A V I Y H I S - R C I L
b S E R A A N I L Q - S T T S A S V A Y -
c Q N E P P T S C S N L P H Q L A L H I K

AATGGACAGTTGAAGCATGAGTCCATAGCATTAATTTTACATGCCAACGGTTTTT
-----+-----+-----+-----+-----+ 8040
TTACCTGTCAACTTCTCGTACTCAAGGTATCGTAATAAAATGTACCGTTGCCCAAAAA

7981

a N G Q L K S M S I A L F L H G Q R V F
b M D S - R A - V P - H Y F Y M A N G F F
c W T V E E H E F H S I I F T W P T G F L

TGCGCAATATAGTCGGAGAGCTTTCTGCTGAGCGACAACAGATTATCCTGCCCTCCTG
-----+-----+-----+-----+-----+ 8100
ACGCGTTATATCAGCCTCTCGAAAGACGACTCGCTGTTGTCTAAATAGGACGGGAGGAC

8041

a C A I - S E S F L L S D N R F I L P L L
b A Q Y S R R A F C - A T T D L S C P S C
c R N I V G E L S A E R Q Q I Y P A P P V

Figure 11AS

TGGTAGTCCCTGTATATTACAGGCTGGTGCCAGCTTACATTAATCGAACTTGAGTCTATCG 8160

8101

ACCATCAGGACATATAAGTCCGACACCGTGCGAATGTAATTAGCTTGAACCTCAGATAGC

a W - S L Y I Q A G A S L H - S N L S L S
b G S P C I F R L V P A Y I N R T - V Y R
c V V P V Y S G W C Q L T L I E L E S I E

AAATCGGCATGGCGTTTCGGATTCAATGCTTCGGCGACATCAGACTCGGTTTTTGGCTA 8220

8161

TTTAGCCGTACCCGCAAGCCTAAGTAACGAAGCCGCTGTAGTCTGAGCCAAAACGAT

a K S A W A F G F I A S A T S D S V F L L
b N R H G R S D S L L R R H Q T R F F C Y
c I G M G V R I H C F G D I R L G F F A I

TTCAACTACCTGGGGAATCTACGCAAGGGTGTGCTGACAGAGGATAACACGATGAAAT 8280

8221

AAGTTGATGGACCCCTTAGATGCGTTCACACGACTGTCTCCTATTGTGCTACTTTA

a F N Y L G E S T Q G C C - Q R I T R - N
b S T T W G N L R K G V A D R G - H D E I
c Q L P G G I Y A R V L L T E D N T M K F

Figure 11AT

TTGACGAAATTAGTCCAGGATATCGAAACGCTACTTGGTTCAGGGAGCCCAATGTCAAAGA 8340

8281

AACTGCTTAATCAGGTCCTATAGCTTTGCGATGAACGAGTCCCTCGGGTTACAGTTTCT

a L T N - S R I S K R Y L R Q G A Q C Q R
b - R I S P G Y R N A T C V R E P N V K E
c D E L V Q D I E T L L A S G S P M S K S

GTGACGGAACGTCTTCAGTCGAACTTGAGCAGATACCACAACAGGTGCTCTTTGAGGTCTG 8400

8341

CACTGCCTTGCAGAAAGTCAGCTTGAACTCGTCTATGGTGTGTCCACGAGAAACTCCAGC

a V T E R L Q S N L S R Y H N R C S L R S
b - R N V F S R T - A D T T T G A L - G R
c D G T S S V E L E Q I P Q Q V L F E V G

GACGTGCGAGTCTGGAAATTGGACAATTACGACAACCTTAAACGGGGACGTTTGCCTG 8460

8401

CTGCACGCTCAGACCTTTAACCTGTTAATGCTGTGAATTTTGCCCCCTGC AAAACGGAC

a D V R V W K L D N Y D N L K R G T F C L
b T C E S G N W T I T T T - N G G R F A C
c R A S L E I G Q L R Q L K T G D V L P V

Figure 11AU


```

8461 TAGTGATGTTTTCGCCAGAGGTGACGATAAGAGTAATGACCGTATTATTGGCAAG + 8520
-----+-----+-----+-----+-----+-----+-----+
ATCCACCTACAAAACGGGTCTCCACTGCTATTCTCATTTACTGGCATAATAACCGTTC

a - V D V L R Q R - R - E - M T V L L G K
b R W M F C A R G D D K S K - P Y Y W A R
c G G C F A P E V T I R V N D R I I G Q G

8521 GTGAGTTGATTGCCTGTGGCAATGAATTTATGTGCGTATTACACGTTGGTATCTTTGCA + 8580
-----+-----+-----+-----+-----+-----+
CACTCAACTAACGGACACCGTTACTTAAATACCACGCATAATGTGCAACCATAGAAACGT

a V S - L P V A M N L W C V L H V G I F A
b - V D C L W Q - I Y G A Y Y T L V S L Q
c E L I A C G N E F M V R I T R W Y L C K

8581 AAAATACAGCGTAAACCTGATAAGAAAAATAATATGCGAACAAATAATAGGTTCCAGG + 8640
-----+-----+-----+-----+-----+-----+
TTTATGTCGCAATTGGACTATTCTTTTATTATACGCTTGTATTATTCGCAAGGTCC

end yscQ*
a K I Q R K P D K K N N M R T I - - R S R
b K Y S V N L I R K I I C E Q Y N S V P G
c N T A - T - - E K - Y A N N I I A F Q V

```

Figure 11AV

TCGTGTCATGAGAGATACAGTATGTCTTTACCCGATTGCGCTTTGCAACTGATTGGTATA + 8700
-----+-----+-----+-----+-----+-----+-----+-----+
AGCACAGTACTCTCTATGTATACAGAAATGGGCTAAGCGGAAACGTTGACTAACCATAT

8641

start yscr*?

a S C H E R Y S M S L P D S P L Q L I G I
b R V M R D T V C L Y P I R L C N - L V Y
c V S - E I Q Y V F T R F A F A T D W Y I

TTGTTTCGCTTCAATACTGCCTCTCATTATCGTCATGGGAACCTTTCCTTAAACTG + 8760
-----+-----+-----+-----+-----+-----+-----+
AACAAAGACGAAAGTTATGACGGAGAGTAATAGCAGTACCCCTTGAAGAAAGGAATTGAC

8701

a L F L L S I L P L I I V M G T S F L K L
b C F C F Q Y C L S L S S W E L L S L N W
c V S A F N T A S H Y R H G N F F P - T G

GCGGTGGTATTTTCGATTTTACGAAATGCTCTGGGTATTCAACAAGTCCCCCAAATATC + 8820
-----+-----+-----+-----+-----+-----+-----+
CGCCACCATAAAGCTAAATGCTTTACGAGACCCATAGTTGTTTCAGGGGGTTTATAG

8761

a A V V F S I L R N A L G I Q Q V P P N I
b R W Y F R F Y E M L W V F N K S P Q I S
c G G I F D F T K C S G Y S T S P P K Y R

Figure 11AW

8821 GCACGTATGGCCTTGGCCTTGACTTTCCTTATTCAATATGGGCGCGCTATTAGCT 8880
-----+-----+-----+-----+-----+-----+-----+
CGTGACATACCGGAACGGAACATGAAGGAATAAGTAATACCCGGCTGCGATAATCGA

a A L Y G L A L V L S L F I M G P T L L A
b H C M A L R L Y F P Y S L W G R R Y - L
c T V W P C A C T F L I H Y G A D A I S C

8881 GTAAAGAGCGCTGGCATCCGGTTCAGGTCGCTGGCGCTCCTTCTGGACGCTGAGTGG 8940
-----+-----+-----+-----+-----+-----+-----+
CATTTCTCGGACCGTAGGCCAAGTCCAGCGACCGGAGGAAGACCTGCAGACTCACC

a V K E R W H P V Q V A G A P F W T S E W
b - K S A G I R F R S L A L L S G R L S G
c K R A L A S G S G R W R S F L D V - V G

8941 GACAGTAAAGCATTAGCGCCTTATCGACAGTTTTCGCAAAAAAACTCTGAAGAGAAGGAA 9000
-----+-----+-----+-----+-----+-----+-----+
CTGTCATTTCGTAATCGCGGAATAGCTGTCAAAAACGTTTTTTTGAGACTTCTCTTCCTT

a D S K A L A P Y R Q F L Q K N S E E K E
b T V K H - R L I D S F C K K T L K R R K
c Q - S I S A L S T V F A K K L - R E G S

Figure 11AX

Figure 11AY

9181 ATACTGCTGGCTATGGGATGATGATGGTGTGCGCGATGACCATTTTCATTACCGTTTAAG 9240
-----+-----+-----+-----+-----+-----+-----+
TATGACGACCGATACCCCTACTACTACACAGCGGCTACTGGTAAAGTAATGGCAAATTC

a I L L A M G M M M V S P M T I S L P F K
b Y C W L W G - - W C R R - P F H Y R L S
c T A G Y G D D D G V A D D H F I T V - A

9241 CTGCTAATATTTTACTGGCAGCGGTTGGATCTGACACTGGCGCAATTGGTACAGAGC 9300
-----+-----+-----+-----+-----+-----+-----+
GACGATTATAAAATGACCGTCCGCCCAACCCTAGACTGTGACCGCGTTAACCATGTCTCG

a L L I F L L A G G W D L T L A Q L V O S
b C - Y F Y W Q A V G I - H W R N W Y R A
c A N I F T G R R L G S D T G A I G T E L
end yscR*

9301 TTTTCATGAATGATCTGAATTGACGCAATTGTAAACGCACTTTTATGGATCGTCCTTT 9360
-----+-----+-----+-----+-----+-----+-----+
AAAAGTACTTAAAGACTTAAGTAACTGCGTTAAACATTGCGTTGAAAAATACCTAGCAGGAAA

start yscS*
a F S - M I L N - R N L - R N F Y G S S F
b F H E - F - I D A I C N A T F M D R P F
c F M N D S E L T Q F V T Q L L W I V L F

Figure 11AZ

9361

9420

ந டு ப

9421

ਸੁ ਧੁ ਧੁ

9487

ਭ ਮੁ ਪ

Figure 11BA

AGATAATGTTACGAATTGGAGAGCATGGTTGAATGGCACACAGGTAATGAGTGGCTTA 9600

9541

TCTATTACAATGCTTAACCTCTCGTACCAACTTACCGTGTGTCCATTACTCACCGAAT

end yscS* start yscT*

a R - C Y E L E S M V E W H N R - M S G L
b D N V T N W R A W L N G T T G K - V A Y
c I M L R I G E H G - M A O Q V N E W L I

TTGCATTGGCTGTGGCTTTTATTTCGACCATGAGCCTTCTTTATTACTTCCCTTATTAA 9660

9601

AACGTAACCGACACCGAAATAAGCTGGTAACCTCGGAAAGAAATAATGAAGGGAATAATT

a L H W L W L L F D H - A F L Y Y F P Y -
b C I G C G F Y S T I E P F F I T S L I K
c A L A V A F I R P L S L S L L L P L L K

AAAGTGGCAGTTAGGGCCGCACCTTTACGTAATGGCGTGTATGTCACTTACCTTTC 9720

9661

TTTCACCGTCAAAATCCCCGGCGTGAAATGCATTACCGCACGAATACAGTGAATGGAAAG

a K V A V - G P H F Y V M A C L C H L P F
b K W Q F R G R T F T - W R A Y V T Y L S
c S G S L G A A L L R N G V L M S L T F P

Figure 11BB

9721 CGATATTACCAATCATTTACCAGCAGAAGATTATGATGCATATTGGTAAAGATTACAGTT
-----+-----+-----+-----+-----+-----+-----+
GCTATAATGGTTAGTAAATGGTCGTCCTTCTAATACTACGTATAACCATTTCTAATGTCAA

a R Y Y Q S F T S R R L - C I L V K I T V
b D I T N H L P A E D Y D A Y W - R L Q L
c I L P I I Y Q Q K I M M H I G K D Y S W

Tn insertion P9B7



9781 GGTTAGGGTTAGTCACTGGAGAGGTGATTATTGGTTTTCATTTGGTTTGTGCGGCGG
-----+-----+-----+-----+-----+-----+-----+
CCAATCCCAATCAGTGACCTCTCCACTAATAACCAAAAAGTTAACCCAAAACACGCCGCC

a G - G - S L E R - L L V F Q L G F V R R
b V R V S H W R G D Y W F F N W V L C G G
c L G L V T G E V I I G F S I G F C A A V

9841 TTCCCTTTGGGCCGTTGATATGGCGGGTTTCTGCTTGATACTTTACGTGGCGGACAA
-----+-----+-----+-----+-----+-----+-----+
AAGGAAACCCGGCAACTATACCGCCCCCAAGACGAACTATGAATGCACCGCGCTGTT

a F P F G P L I W R G F C L I L Y V A R Q
b S L L G R - Y G G V S A - Y F T W R D N
c P F W A V D M A G F L L D T L R G A T M

Figure 11BC

TGGGTACGATATTCAATTCTACAATAGAAGCTGAACCTCAGTTTGGCTTGCTTTTCA 9960

9901

ACCCATGCTATAAGTTAAGATGTTATCTTCGACTTTGGAGTGAAAAACCGAACGAAAAGT

a W V R Y S I L Q - K L K P H F L A C F S
 b G Y D I Q F Y N R S - N L T F W L A F Q
 c G T I F N S T I E A E T S L F G L L F S

GCCAGTTCTGTGTGTTATTTCTTTATAAGCGGGCATGGAGTTTATATAAACATTC 10020

9961

CGGTCAAGAACACACAATAAAGAAATATTCGCCCGGTACCTCAAATATAATTGTAAG

a A S S C V L F S L - A A A W S L Y - T F
 b P V L V C Y F L Y K R R H G V Y I K H S
 c Q F L C V I F I S G G M E F I L N I L

TGTATGAGTCATATCAATATTTACCACCGGGCGTACTTTATTTATTTGACCAGCAATTT 10080

10021

ACATACTAGTATAGTTATAAATGGTGTCCCGCATGAATAATAAACTGGTCGTTAAAA

a C M S H I N I Y H Q G V L Y Y L T S N F
 b V - V I S I F T T R A Y F I I - P A I F
 c Y E S Y Q Y L P P G R T L L F D Q Q F L

Figure 11BD

```

10081 TAAAATATATCCAGCAGAGTGGAGAACGCTTTATCAATTATGTATCAGCTTCTCTTTC 10140
-----+-----+-----+-----+-----+-----+-----+
ATTTTATAGGTCGCTCACCCTCTTGCGAAATAGTTAATACATAGTCGAAGAGAGAAG

a - N I S R Q S G E R F I N Y V S A S L F
b K I Y P G R V E N A L S I M Y Q L L S S
c K Y I Q A E W R T L Y Q L C I S F S L P

10141 CTGCCATAATATGATGGTATTAGCCGATCTGGCTTTAGGTCCTTTTAAATCGGTCGGCAC 10200
-----+-----+-----+-----+-----+-----+-----+
GACGGTATTATACATACCATAATCGGCTAGACCGAAATCCAGAAAATTTAGCCAGCCGTG

a L P - Y V W Y - P I W L - V F - I G R H
b C H N M Y G I S R S G F R S F K S V G T
c A I I C M V L A D L A L G L L N R S A Q

10201 AACAAATTGAATGTGTTTTTCTCTCAATGCCGCTCAAAAGTATATTGGTTCTACTGACGY 10260
-----+-----+-----+-----+-----+-----+-----+
TTGTTAACTTACACAAAAGAGAGTTACGGCGAGTTTTCATATAACCAAGATGACTGCR

a N N - M C F S S Q C R S K V Y W F Y - X
b T I E C V F L L N A A Q K Y I G S T D X
c Q L N V F F S M P L K S I L V L L T X

```

Figure 11BE

10261 CCTGATCTCATCCCTTATGCTCTTCATCACTATTGTTGAAAGCGATAAATTTATAT + 10320
 -----+-----+-----+-----+-----+-----+-----+-----+
 GGACTAGAGTAAGGAATACGAGAAGTAGTGATAAACCAACTTTCGCTATTTTAAATATA

 a P D L I P L C S S S L F G - K R - I L Y
 b L I S F P Y A L H H Y L V E S D K F Y I
 c - S H S L M L F I T I W L K A I N F I F

 10321 TTATCTAAAGACTGGTTTCCATCTGTATGAGCGAGAAAACAGAACGCTACAGAAAAG + 10380
 -----+-----+-----+-----+-----+-----+-----+
 AATAGATTTTCTGACCAAGGTAGACATACTCGCTCTTTTGTCTTGTGCGGATGCTTTTC

 end yscI* start yscU*
 a L S K R L V S I C M S E K T E Q P T E K
 b Y L K D W F P S V - A R K Q N S L Q K R
 c I - K T G F F H L Y E R E N R T A Y R K E

 10381 AAATTACGTGATGCCGTAAGGAAGGCAGGTTGTCAAAAAGTATTGAAATAACATCATTA + 10440
 -----+-----+-----+-----+-----+-----+-----+
 TTTAATGCACCTACCGCATTCCTTCCCGTCCACAGTTTTCATAACTTTATTGTAGTAAT

 a K L R D G R K E G Q V V K S I E I T S L
 b N Y V M A V R K G R L S K V L K - H H Y
 c I T - W P - G R A G C Q K Y - N N I I I

Figure 11BF

TTTCAGCTGATTGCGCTTTATTGTATTTTCATTCTTACTGAAAGATGATTTTGATA
 -----+-----+-----+-----+-----+-----+-----+
 AAAGTCGACTAACGCGAAATAAACATAAAGTAAAGAAATGACTTTTCTACTAAACTAT

F Q L I A L Y L Y F F H F F T E K M I L I
F S - L R F I C I F I S L L K R - F - Y
S A D C A L F V F S F L Y - K D D F D

CTGATTGAGTCAATAACTTTTCCATTACATTACAATTAGTAAATAAAACCATTCTTATGCATTA
-----+-----+-----+-----+-----+-----+-----+-----+
GACTAACTCAGTTATTGAAAGTGTAATGTTAATCATTTATTGTTGTTAAAGAAATACGTAAT

L I E S I T F T L Q L V N K P F S Y A L
 - L S Q - L S H Y N - - I N H F L M H -
 D - V N N F H I T I S K - T I F L C I

ACGCAATTGAGTCATGCTTTAATAGAGTCACTGACTTCTGCACTGCTGTTTCTGGCGCT
-----+-----+-----+-----+-----+-----+-----+
TGCGTTAACTCAGTACGAAATTATCTCAGTGACTGAAGACGTGACGACAAAGACCCGCGA

T Q L S H A L I E S L T S A L L F L G A L
R N - V M L - - S H - L L H C C F W A L
A I E S C F N R V T D F C T A V S G R

Figure 11BG

GGGGTAATAGTTGCTACTGTGGTAGCGTGTTCCTTCAGGTGGGGTGGTTATTGCCAGC
-----+-----+-----+-----+-----+ 10680
CCCCATTATCAACGATGACACCCCATCGCACAAAGAAGTCCACCCCAACCAATAACGGTCG

10621

a G V I V A T V G S V F L Q V G V V I A S
b G - - L L L W V A C F F R W G W L L P A
c G N S C Y C G - R V S S G G G Y C Q Q

AAGGCCATTGGTTTAAAGCGAGCATATAAATCCGGTAAGTAATTTAAGCAGATATTC
-----+-----+-----+-----+ 10740
TTCCGGTAACCAAAATTTTCGCTCGTATATTAGGCCATTCAATAAATTCGTCATAAG

10681

a K A I G F K S E H I N P V S N F K Q I F
b R P L V L K A S I - I R - V I L S R Y S
c G H W F - K R A Y K S G K - F - A D I L

TCTTTACATAGCGTAGAATTATGTAAATCCAGCCCTAAAAGTTATCATGCTATCTCTT
-----+-----+-----+-----+ 10800
AGAAATGTATCGCATCATCTTAATACATTTAGTCCGATTTTCAATAGTACGATAGAGAA

10741

a S L H S V V E L C K S S L K V I M L S L
b L Y I A - - N Y V N P A - K L S C Y L L
c F T - R S R I M - I Q P K S Y H A I S Y

Figure 11BH

ATCTTTGCCCTTTTCTTTTATTATTATGCCAGTACTTTTCGGGCGCTACCGTACTGTGGG
-----+-----+-----+-----+-----+ 10860
TAGAAACGGAAAAAGAAATAATAACGGTTCATGAAAAAGCCCGCATGGCATGACACCC

10801

a I F A F F Y Y A S T F R A L P Y C G
b S L P F S F I I M P V L F G R Y R T V G
c L C L F L L L C Q Y F S G A T V L W V

TTAGCCTGTGGCGTCTTGTGGTTTCTTTTAAATAAAATGGTTATGGGTAGGGGTGATG
-----+-----+-----+-----+-----+ 10920
AATCGGACACCGCACGAACACCAAGAAAGAAATTTTACCAATACCCATCCCACTAC

10861

a L A C G V L V V S S L I K W L W V G V M
b - P V A C L W F L L - - N G Y G - G - W
c S L W R A C G F F F N K M V M G R G D G

GTTTTTATATCGTCTGGCATACTGGACTATTCTTTTCAATATTATAAGATTAGAAAA
-----+-----+-----+-----+-----+ 10980
CAAAAAATATAGCAGCAACCGTATGACCTGATAAGAAAGTTATAATATTCTAATCTTTT

10921

a V F Y I V V G I L D Y S F Q Y Y K I R K
b F F I S S L A Y W T I L F N I I R L E K
c F L Y R R W H T G L F F S I L - D - K S

Figure 11BI

GCTATCTAAAATGAGTAAAGATGACGTAAACAGGAGCATAAAGATCTGGAGGGCGGACC
-----+-----+-----+-----+-----+-----+ 11040
CGATAGATTTTACTCATTTCTACTGCATTTTGTCCCTCGTATTTCTAGACCTCCCGCTGG

10981

a A I - K - V K M T - N R S I K I W R A T
b L S K N E - R - R K T G A - R S G G R P
c Y L K M S K D D V K Q E H K D L E G D P

Tn insertion P12F5



CTCAAATGAAGACGGCGGTCGGAAATGCAGAGTGAAATACAAAGTGGAGTTAGCTCA
-----+-----+-----+-----+-----+-----+ 11100
GAGTTACTTCTGCGCGCAGCCCTTACGTCTCACTTTATGTTTCACCCCTCAAATCGAGT

11041

a L K - R R G V G N A E - N T K W E F S S
b S N E D A A S E M Q S E I Q S G S L A Q
c Q M K T R R R K C R V K Y K V G V - L N

ATCTGTTAAACAATCTGTTGCGGTAGTGCCTAATCCACGCATATTGCGGTTTGTCTTGG
-----+-----+-----+-----+-----+-----+ 11160
TAGACAATTTGTTAGACAACGCCCATCAGCATTAGGTTGCGTATAACGCCAAACAGAACC

11101

a I C - T I C C G S A - S N A Y C G L S W
b S V K Q S V A V R N P T H I A V C L G
c L L N N L L R - C V I Q R I L R F V L A

Figure 11BJ

11161

a L S S H R Y A N T T R P G K R Q - C S S
b Y H P T D M P I P R V L E K G S D A Q A
c I I P P I C Q Y H A S W K K A V M L K L

11221

a - L Y C - H R - T Q L H P R C - K C - A
b N Y I V N I A E R N C I P V E N V E L
c T I L L T S L N A T A S P L L K M L S W

11281

a G P L I I F - S G T R R - N S - N V I -
b A R S L F F E V E R G D K I P E T L F E
c P A H Y F L K W N A E I K F L K R Y L N

Figure 11BK

11341 ACCCGTTGCAGCCTTGTTACGTATGGTGATGAAGATAGATTATGCGCATCTACCGAAAC 11400
-----+-----+-----+-----+-----+-----+-----+
TGGGCAACGTCGGAACAATGCATACCACCTACTTCTATCTAATAACGCGTAAGATGGCTTTG

a T R C S L V T Y G D E D R L C A F Y R N
b P V A A L L R M V M K I D Y A H S T E T
c P L Q P C Y V W - - R - I M R I L P K H
end yscU*

11401 ACCATAAATGCTTTTGGTATGCTTCTTCAGGCCACTGCGAAGGTTAAGAGGGTAATAGCG
-----+-----+-----+-----+-----+-----+-----+
TGGTATTACGAAACCATACGAAGAAGTCCGGTGACGCTTCCAATTCTCCCATATATCGC

a T I N A F G M L L Q A T A K V K R V I A
b P - M L L V C F F R P L R R L R G - - R
c H K C F W Y A S S G H C E G - E G N S V

11461 TATAGAGCAGTGCTTGACGATAAAGGTGAGAGACTGAAAATAATCGCTTTAGCCTGGCA 11520
-----+-----+-----+-----+-----+-----+-----+
ATATCTGTCACGAACTGCTATTTCCTACTCTCTGACTTTTATTATAGCGAAAATCGGACCGT

a Y R A V L D D K G E R L K I I A F S L A
b I E Q C L T I K V R D - K - S L L A W H
c - S S A - R - R - E T E N N R F - P G T

Figure 11BL

CAAGCACCAGATAGCGTATTATAAAATTAACAAGATAATGGATTGGTGCGTCTGAATGG 11580
-----+-----+-----+-----+-----+-----+-----+-----+
GTTGCTGGTCTATCGCATAATAATTTAAATTTGTTCTATTACCTAACACGACGACTTACC

11521

a Q A P D S V L - N - T R - W I G A S E W
b K H Q I A Y Y K I K Q D N G L V R L N G
c S T R - R I I K L N K I M D W C V - M D

ACTCGAACCACTCGACCCCAACCATGTCAAGGTGGTGCTCTAACCAACTGAGCTATGAAC 11640
-----+-----+-----+-----+-----+-----+-----+-----+
TGAGCTTGGTGAGCTGGGGGTGACAGTTCCACCACGAGATTGGTTGACTCGATACTTG

11581

a T R T T R P P P C Q G G A L T N - A M N
b L E P L D P H H V K V V L - P T E L - T
c S N H S T P T M S R W C S N Q L S Y E R

GGCAACGTTGTAGTGACAACGGGGACGAATATTAGCGTCACAACCGCAATGAGGCAAGA 11700
-----+-----+-----+-----+-----+-----+-----+-----+
CCGTGCAACATCCACTGTTGCCCTGCTTATAATCGCAGTGTGGCGTTACTCCGTTCT

11641

a G N V V G D N G D E Y - R H N R N E A R
b A T L - V T T G T N I S V T T A M R Q E
c Q R C R - Q R G R I L A S Q P Q - G K R

Figure 11BM

11701 GGGAAATCGCAATTTCTTCTGAAATCACCTGATTGCGGTGGAATATGCAACATGTCG 11760
-----+-----+-----+-----+-----+-----+-----+
CCCTTTAGCGTTAAAGAAGGACTTTAGTGAGCTAACGCCACCTTTATACGTTGTACAGC

a G K S Q F S S - N H L I A V E I C N M S
b G N R N F L P E I T - L R W K Y A T C R
c E I A I F F L K S P D C G G N M Q H V E

11761 AGAAATAGCCGCCATGCGACGGCTATCGTCGTATTATCGGAGCGCTGCAAAATGATG 11820
-----+-----+-----+-----+-----+-----+-----+
TCTTTATCGGCGTACGCTGCCGATAGCAGCATAATAGCCTCGCGGACGTTTACTAC

a R K - P P C D G Y R I I G A R C K M M
b E N S R H A T A I V V L S E R A A K - W
c K I A A M R R L S S Y Y R S A L Q N D G

11821 GCGGACGGCTGACGTTGTAGATAGCGCATCCGTAGCATCATTAACACCCGCCGAGGTC 11880
-----+-----+-----+-----+-----+-----+-----+
CGCCTGCCGACTGCAACATCTATCGGTAGGCATCGTAGTAATTGTGCGGCGGCTCCAG

a A D G - R C R - R I R S I I N T A A E V
b R T A D V V D S A S V A S L T P P P R S
c G R L T L - I A H P - H H - H R R R G Q

Figure 11BN

Confirmation of 09/20/95
RPM 10100 (3)

11881 AGGCCGATGAACCCCATCCAGAAGCCTGCCGTCCCATACGATCCACCAATCC + 11940
-----+-----+-----+-----+-----+-----+-----+
TCCGGCTACTTGGGTAGGTCTTCGACGGCCAGGTATGCTAGGTGGTGTAGG

a R P M M N P I Q K P A G P I R S T T K S
b G R - - T P S R S L P V P Y D P P P N P
c A D D E P H P E A C R S H T I H H Q I R

11941 GTTAACGCCAGGATATAACCGCTGGTAAACCTAACACCCAGTAGCGGTAAAGGTGATA + 12000
-----+-----+-----+-----+-----+-----+-----+
CAATTGCGTCCCTATATTGGCGACCCATTGGATTGTGGTCATCCGCCATTTCCTACTAT

a V N A R I - P L G K P N T Q - A V K V I
b L T P G Y N R W V N L T P S R R - R - -
c - R Q D I T A G - T - H P V G G K G D K

12001 AAAAGATGGAACGGTATCTTTATAACCGCGCAGAAATACCGTGCCGATAACCTGTATA + 12060
-----+-----+-----+-----+-----+-----+-----+
TTTTTCTACCTTGGCATAGAAATATTGGCGGCTCTTATGGCGACGGCTATTGGACATAT

a K K M E R V S L - P R R I P L P I T C I
b K R W N A Y L Y N R A E Y R C R - P V -
c K D G T R I F I T A Q N T A A D N L Y R

Figure 11BO

12061

അ മു ഹ

12121

π ρ υ

12181

ന മ ഉ

Figure 11BP

a N T - W R N Q Q R N D R K - R H F K E Q
b I P N G E T S S A T T A N N V T S K N S
c Y L M A K P A A Q R P Q I T S L Q R T A

Figure 11BQ

Figure 11BR

Continuation of 09/201,945
R 275 101 Con (3)

GTTCCGGCATACCAAAATGGCCATAGATAAAATATAGTTCACCGGAATATTCACCAGCA 12660

12601

CAAGGCCGTATGGTTTACCGGTATCTATTTTATATCAAGTGGCCTTATAAGTGGTCGT

a V P A Y Q N G H R - K Y S S P E Y S P A
b F R H T K M A I D K N I V H R N I H Q Q
c S G I P K W P - I K I - F T G I F T S R

GGCCAAAATCCCATCACCATACCCGGTTTGGTTTGGCCAGACCTTCGCACTGGTTTC 12720

12661

CCGGGTTTATAGGTAGTGGTATGGCCAAACCAAAACCGGTCTGGAAGCGTGACCAAAG

a G P K I P S P Y P V W F W P D L R T G F
b A Q K S H H T R F G F G Q T F A L V S
c P K N P I T I P G L V L A R P S H W F R

GCGTACCTGAAAGAAAGGTATCCTGCGCCCCACAGCAGCGCGGAAGATAACCCACGG 12780

12721

CGCGATGGACTTCTTTCCATAGGACGCGGGGTGTCGTCGCGGCTTCTATTGGGTGCC

a A L P E R K G I L R P T A A R E D N P R
b R Y L K E K V S C A P Q Q R A K I T H G
c A T - K K R Y P A P H S S A R R - P T A

Figure 11BS

12781

॥ ३ ॥

12841

ந டு ப

12901

അ മു പ

Figure 11BT

Figure 11BU

13141
TGGTATACTTCTGCACGTATTCACCTTTTATTTTGTTGTATATGAAAGACTAAAGCC
-----+-----+-----+-----+-----+-----+ 13200
ACCATATGAAGACGTGCATAAGTGGAAAAATAAAACAACAATATACTTTCTGATTTTTCGG

a W Y T S A R I H L L F C C Y M K D - K A P

b G I L L H V F T F Y F V V I - K T K K P

c V Y F C T Y S P F I L L L Y E R L K S R

13201
GCCGAAGTGGCAGCCAAAAGAAATAGCAGGGGAAATTCAGTCTATTGTAGCGGGTATT
-----+-----+-----+-----+-----+-----+-----+ 13260
CGGCTTACCGTCGGTTTCTTTATCGTCCCCCTTTAAAGTCAGATAACATCGCCCCATAA

a A E V A A K R N S R G N F S L L - R G I
b P K W Q P K E I A G E I S V Y C S G V L
c R S G S Q K K - Q G K F Q S I V A G Y Y

ACTATTTCTCCAGTGAAAAAACAGTTGTTAACGGCGCATTTGCTGGCAAGCTGTTTTCCTCA
 13261
 -----+-----+-----+-----+-----+-----+-----+-----+
 TGATAAAGAGGTCACCTTTTGTGTCACAATGCCCCGTAAACGACCGTTTCGACAAAAAGGT
 13320

a	T	I	S	P	V	K	K	Q	L	L	T	A	H	C	W	Q	A	V	F	F
b	L	F	L	Q	-	K	N	S	C	-	R	R	I	A	G	K	L	F	F	H
c	Y	F	S	S	E	K	T	V	V	N	G	A	L	L	A	S	C	F	S	T

Figure 11Bv

13321 CCTGCTATTGTGCTGAACAGTTCTGCTTTTATTATTATTTTTCAGGAGTTGAAGATATGTTTAC 13380
-----+-----+-----+-----+-----+-----+-----+
GGACGATAACACGACTTGTCAAGACGAAAAATAAAGTCCTCAACTTCTATACAAATG

a P A I V L N S S A F I Y F R S - R Y V Y
b L L L C - T V L L L F I S G V E D M F T
c C Y C A E Q F C F Y L F Q E L K I C L R

13381 GGGGATCGTACAGGGTACCGCGAAACTGGTATCGATA 13417
-----+-----+-----+-----+-----+-----+
CCCCTAGCATGTCCCATGGCGCTTTGACCATAGCTAT

a G D R T G Y R E T G I D
b G I V Q G T A K L V S I
c G S Y R V P R N W Y R

Figure 11BW

Figure 12A
DNA Sequence of VGC II cluster C

Tn insertion P9B4
↓
GGATCCTTTTCTTAATGCTGCTAACGTTTCTTGCAAAATGCGTTGATGAGATTCATCC 60
1 - - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +
CCTAGGAAAAGAAATTACGACGATTGCAAGAACGTTTACGCAACTACTCTAAGTAGG
AGTACACCACCTGATAACAAAGAGCGNCGCATTTGGCNWAMMWTKRNNMRNNSCNNNACTA 120
61 - - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +
TCATGTGTGACTATTGTTTTCTCGCNGCGTAACCGNWTKKWAMYNNKYNNSGNNNTGAT
Tn insertion P7A3
↓
AACCGTTCTCTATTATCGCAGAAATAATATCATCCCCCTGAGACTGATGAGAGTGACTAA 180
121 - - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +
TTGGCAAGAGATAATAGCGTCTTTATTATAGTAGGGGACTCTGACTACTCTCACTGATT
TCTGCCAGTGCAATAACCCGGGAATATCTGCAAGTAATGTTGAACCTTGCGCCATTGCT 240
181 - - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +
AGACGGTCACGTTATTGGGCCCTTATAGACGTTCAATACCAACTTGGAAACGCGGTAACGA
Tn insertion P964
↓
GATCCATTGTATATCATGAATTAACACGCTCCCCGGCCCTTCGCTGGATACTCAG 300
241 - - - - - + - - - - - + - - - - - + - - - - - + - - - - - + - - - - - +
CTAGGTAACATATAGTAGTACTTAATTGTGCGAGGGCGGGAAGCGACCTATGAAGTC

Figure 12B

301 CATNSSGGTAACCCATTTTATCAAAACATCCTGCACCTTCTCGTACCAATAAGTCATCAC
-----+-----+-----+-----+-----+-----+
GTANSSCCATTGGGTAAATAAGTTTGTAGGACGTGAAGAGCATGGTTATTCAGTAGTG 360

361 AGATTACACCATCCCGATACATGACCCCCCATGATTCGAGAGTCGCTCTCACCCTTTGGCA
-----+-----+-----+-----+-----+-----+
TCTAATGTGGTAGGGCTATGTACTGGGGGTACTAAGCTCTCAGCGAGAGTGGAAAAACGT 420

421 TCTGTTCGCTTGACGAGCAATAACCGACAACCTGCAGGCTGCCATCTTCTTCCATTGCG
-----+-----+-----+-----+-----+-----+
AGACAAGCGAACTGCTCGTTATTGGCCCTGTTGACGTCGACGGTAGAAGAAAGGTAACGC 480

481 CCCGCACATAATGAATATTGCTTTTGTCTAATAAAACTTAACCCGCAAGGTAAGTCAT
-----+-----+-----+-----+-----+-----+
GGCGGTGTATTACTTATAACGAAAACAGATTATTTTGAATTGGCGGTTTCCATTTCAGTA 540

541 TTACCGTTTCAGGCTGACCACTAATACTTAACAGGACACCCCATTCACCCGATGAAAATCA
-----+-----+-----+-----+-----+-----+
AATGGCAAAGTCCGACTGGTGATTATGAATTGTCTGTGGGTAAGGTGGCTACTTTTAGT 600

601 AGAATACGCCAGCCAAACCACAGTACCCCTGATCTGGAACGGGTATTGATAATCAGCAA
-----+-----+-----+-----+-----+-----+
TCTTATGCGGTCGGTTGGTGCATGGGACTAGACCTTTGCCCATAAACTATTAGTCGTT 660

Continuation of 09/201,945
R 2015 10/Con(3)

Figure 12C

661 GTTCACAATCCTGTTTACCAACGCGATASSCACTCCCGCAACCTGCAAAACCCACTGG 720
-----+-----+-----+-----+-----+-----+
CAAGTGTAGGACAAATGGTTTGGCGCTATSSGTGAGGCGGTGACGCTTTTGGGTGACC
721 ATGGTAGCGGCTTATTTGGATTAAATCTGCGGCCATTAACTCTAACTCTGGCTTTCCCGG
-----+-----+-----+-----+-----+-----+
TACCATCGCCGAATAAACCTAATTAGACGCCGGTAATTGAGATTGAGACCGAAAGGGCC
781 CATCAACAAATAACTATCTGCCCTGTTCTCTCAGAAATAATTTTTCATTTATAGCCAGCG
-----+-----+-----+-----+-----+-----+
GTAGTTGTTTATTGATAGACGGACAAGAGAGTCTTATTAAAAAAGTAAATATCGGTCGC
841 AATACAAATATCGCATCCCTTCTCCCCAGTGACAGGTTACCTTCATTCAGCCATACTTC
-----+-----+-----+-----+-----+-----+
TTATGTTTATAGCGTAGGGAAGAGGGGTCACTGTCCAATGGAAGTAAGTCGGTATGAAG
901 CCGGCCCTTGTAACGTGACCTAAAAACGTATTTTCCAGGAACCTTTGGATTAAACCAT
-----+-----+-----+-----+-----+-----+
GGCCGGAACATTTTGCACTGGATTTTTCATATAAAGGTCCTTGAGAAACCTAATTGGTA
961 GAGATATGCCATTATTTACTACTGAGGCTTTAATCAAAAAAGCCTGATTACACTATGTA
-----+-----+-----+-----+-----+-----+
CTCTATACGGTAATAAATGATGACTCCGAAATTAGTTTTTTTTCGGACTAATGTGATACAT 1020

Figure 12D

1021 CTTGAGTCGTATCATTCGGAACAATAATGACCTACAACAGGAATATCGCCCAATAAAGGA 1080
-----+-----+-----+-----+-----+-----+-----+
GAACTCAGCATAGTAACGCTTTGTTTACTGGATGTTGTCCTTATAGCGGTTATTTCCCT
1081 TTTTGTTTTTCGGAGTGGATTTGTTTACCTTGTTTAAACCTCCAGCAATNAGACTTTGC 1140
-----+-----+-----+-----+-----+-----+-----+
AAAACAAAACGCTCACCTAAACAATAATGGAACAATAATTTGGAGGGTCTGTTANTCTGAAACG
1141 CCGGCCAATAATGTGGCTTGCGAANCRAATTTCAGAAATTTTGCACTTCGGGCAGCGGTCT 1200
-----+-----+-----+-----+-----+-----+-----+
GGCCGGTTATTACACCGAACGCTTNGYTAAGTCTTAAACGTGAAGCCCGTCGCCCCAGA
1201 GTNTYGCYTTKGNSTATCACTTTGTTGTCCATCCTGAANTATTAAGATTAAAGCATTATTT 1260
-----+-----+-----+-----+-----+-----+-----+
CANARCGRAAMCNSATAGTGAAACAACAGGTAGGACTTNATAAATTCTAATTTCGTAATAAA
1261 TTTGCGTGCCATTGTCAATTTAAACAAGCGAGGTGTAAACGCGWNAACAAGAACCCGTAGTG 1320
-----+-----+-----+-----+-----+-----+-----+
AAACGACGGTAACAGTAAATTGTTGCTCCACATTGCGCWNNTGTTTCTTGGGCATCAC
1321 ATGGATTCAAGTTTAGCCACTTTTCTCCCTGCAGTTTGGTATAGAAAGTAATATTTTA 1380
-----+-----+-----+-----+-----+-----+-----+
TACCTAAGTTCAAATCGGTGAAAAGAGGACGTCAAAACCATATCTTTTCATTATATAAAAT

Figure 12E

1381 TCCAGCACAGCCTGGATATTATTTAAAGTCACCACAGATGGCTGGAAAGTACATAAGCC 1440
-----+-----+-----+-----+-----+-----+-----+
AGTCGTGTCGGACCTATAATAAATTTCAGTGGTGCTACCGACCCTTTCATGTATTCCGG
1441 TGAGAGCTTTTCCAGGGCATTTCAGACGCACCATAAAGTTTGAGGTATCGCTGATTACC 1500
-----+-----+-----+-----+-----+-----+-----+
ACTCTCGAAAAAGGTCCCGTAAGTCTGCGTGGTATTTCAAACTCCATAGCGACTAATGG
1501 GTTGANNAACCACTAGCACCAACCGTCATTCAAAACCTGTATTGAACGCAATTTTCTTGCCA 1560
-----+-----+-----+-----+-----+-----+-----+
CAACTNNTTGGTGATCGTGGTGGCAGTAAGTTTGGACATAAAGTTGCGTTAAAGAACGGT
1561 CCCAGCGACACTGCCGTTCCCCAGTCGATGCCCTAACTGGTTAATAATCTCCAGCATTAACA 1620
-----+-----+-----+-----+-----+-----+-----+
GGTCCGCTGTGACGGCAAGGGTTCAGCTACGGATTGACCAATTATAGAGGTCGTAATTGT
1621 TCGATAATTTACCGAAATCTCTATCATCTGCTGGCGTTGATCTAATTCTGTGATGAGT 1680
-----+-----+-----+-----+-----+-----+-----+
AGCTATTAAAGTGGCTTTAGAGATAGTAGACGACCGCAACTAGATTAAAGACACTACTCA
1681 TTCCGATACNNNGCCATATTGGNNNCATAATCACGAACGATCACTGCATTTCTGGCGTNGG 1740
-----+-----+-----+-----+-----+-----+-----+
AAGGCTATGNNNCGGTATAACCNNGTATTAGTGCTTGCTAGTGACGTAAAGACCGCANCC

Figure 12F

1741 GTCCGACGAAACATNGGCAATGCCCTGTGTAGGGGTGAACCATTTGTCNTCGATGACGT 1800
 -----+-----+-----+-----+-----+-----+-----+
 CAGCCGTCGTTTGTANCCGTTACGGACACATCGCCCACTTGGTAACAAGNAGCTACTGCA

 1801 CGGGACGCTGTTTACTCATCTCACGCAATACACTAACGACCCCTGNNNAACCACGACG 1860
 -----+-----+-----+-----+-----+-----+-----+
 GCCCTGCGACCAAAATGAGTAGAGTGCCTTATGTGATTGCTGGGACCNNTTGGTGCTGC

 1861 GACTGATCGCGATATTGGTACTGGGTATCCATCGCAGTGGCATACTTAAGCGTGTATATA 1920
 -----+-----+-----+-----+-----+-----+-----+
 CTGACTAGCGCTATAACCATGACCCATAGGTAGCGTCACCGTATGAATTCCGCACATATAT

 1921 CTTACACTCACCGCACTGTCTTTTCGTTTGATTAACGCATTATCCAGCACTGAAGCTAAT 1980
 -----+-----+-----+-----+-----+-----+-----+
 GAATGTGAGTGGCGTGACAGAAAAGCAAACCTAATTGCGTAATAGGTCGTGACTTCGATTA

 1981 TGAATAACGAGTCAGGACGCTGGGAACACCCGCTCACCTCCACAGCTTTGGTACCGGTA 2040
 -----+-----+-----+-----+-----+-----+-----+
 ACTGATTATGCTCAGTCCGTCGACCCCTTGTGGCGAGTGAGGTGTCGAAAACCATGGCCAT

 2041 ATTTCTTTAACCTCGCATCCCGGTGATGAAAGGATATTCTGGCTGCGTAAGTAATGAATG 2100
 -----+-----+-----+-----+-----+-----+-----+
 TAAAGAAATTGGAGCGTAGGGCCACTACTTTCCTATAAGACCGACGCATTCTACTTAC

Tn insertion P7G2

Continuation of 09/201, 945
RPM 5 101 CBL(13)

Figure 12G

2101 AACCGTCCAGTAGATAAAATATTGAAAGTGATAACCTGATGTTTAAATAACGATGCAGGA 2160
-----+-----+-----+-----+-----+-----+
TTGGCAGGTCATCTAATTTATAACTTTCACCTATTGGACTACAAAATTATTGCTACGTCCT
2161 TATACATAAATCATGCTGCCATCAAAACCAGGTAAGCAAATCATATTTGCTGCCAGGTTA 2220
-----+-----+-----+-----+-----+-----+
ATATGTATATTGTACGACGGTAGTTTGGTCCATTTCGTTAGTATAACACGACGGTCCCAAT
2221 TTCAAAATATCGACCGGTGTCAGGCGGAATTTTCCACTAAATGTAGCTGTTATCAAT 2280
-----+-----+-----+-----+-----+-----+
AAGTTTATAGCTGGCCACCAGGTCGCCCTTAAAAAGGTGATTTACATCGACAATAGTTA
2281 GGGCTAATAGTAATAGCCGTATCATAGTTCTCTGAGAGCAGATGTNAAACCTCTGCTAA 2340
-----+-----+-----+-----+-----+-----+
CCCGATTATCATTTATCGGCATAGTATCAAGAGACTCTCGTCTACANTTTTGGAGACGATT
2341 TGGCATTTGTCTGGCATAAAGGTGAAGTCATTACCTTTCCATGATAACTCATCTCTT 2400
-----+-----+-----+-----+-----+-----+
ACCGTAAACAGACCGTATTTCCCACTTCAGTAATGGAAGGTACTATTGAGTAGTGAGAA
2401 TGCTGTATTGAGTATAAATAGTAAATTAAGATTAAACGTTTATTACTACCATTTTATA 2460
-----+-----+-----+-----+-----+-----+
ACGACATAACTCATATTATCATTTTAAATTCATAATTGCAATAAATGATGGTAAATAAT

Continuation of 09/201, 945
RP 101 (Cont.)

2520

2461

2521

2581

2641

Tn insertion P11B9

2701

2761

Figure 12I

2821 WWTAAATGGAATGCCCTTTTAAAACTGCCAGCATGAATCCCTCCTCAGACATAAATGGGAG
 +-----+-----+-----+-----+-----+-----+
 WWAATTACCTTACGGAAATTTTGACGGTCGTACTTAGGGAGGAGTCTGTATTTACCCCTC
 +-----+-----+-----+-----+-----+-----+
2881 TTTCTATCAAAATTCGCTCACAAACCACATCCGTAAAAAGCCCTGATTACATTTATTTCGAC
 +-----+-----+-----+-----+-----+-----+
 AAAGATAGTTTAAGCGAGTGTGGTGTAAGCATTTTTCGGACTAAGTGTAATAAAGCTG
 +-----+-----+-----+-----+-----+-----+
2941 TATACTCTTCTGTACAATATCAGGATGCTGTCTACATATACCTTGTACAGGCGATTCT
 +-----+-----+-----+-----+-----+-----+
 ATATGAGAAGAACATGTTATAGTCCCTACGACAGATGTATATGGAACAGTGTCCGCTAAGA
 +-----+-----+-----+-----+-----+-----+
3001 ATCATTCGGATTTTCCGATAAAATTNMMCAATTACATTTTCAGCATTGACATAAAACTTA
 +-----+-----+-----+-----+-----+-----+
 TAGTAAGCCTAAAAGGCTATTTAANKKGTTAATGTAAAAGTCGTAACGTATTTTGAAT
 +-----+-----+-----+-----+-----+-----+
3061 CAATTTGNAAAATTAATTAATAAACTGTTACGATGTTTTTACATCGCCATCTTATT
 +-----+-----+-----+-----+-----+-----+
 GTTAAACNTTTTAATAAATAATTTATTTGACAAATGCTACAAAATGTAGCGGTAGAAATAA
 +-----+-----+-----+-----+-----+-----+
3121 AAAAAGTAATTGTAGTCATCGACTNGGTTATATATGAAGAAAATTTATCTTCCTAATGATA
 +-----+-----+-----+-----+-----+-----+
 TTTTTCATTAAACATCAGTAGCTGANCCAAATATATACCTCTTAAATAGAAGGATTACTAT
 +-----+-----+-----+-----+-----+-----+

Continuation of 09/201, 945
RPMs 101 Cox(3)

3240

3300

3360

3420

3480

3540

Figure 12K

3541 TCATAACGATATTTCCCTGAGGTGAGCCGGCATCTATCTGTCGGTCCTTCAAAATTGCAC
3600
3541 AGTATTGCTATAAAAGGACTCCACTCGGCCGTAGATAGACAGCCAGGAAGTTTAAACGTG
3600
3601 MGCCGACGCTNAACGGAGAGAACCGTCTCTTTCTGCAGTCCTCTGATATCGATGAAA
3660
3601 KCGGCTGCGANTTGCCCTCTCTTCGTGGCAGAGAAAGACGTCAGGAGACTATAGCTACTTT
3660
3661 ATAGCTTTCGTCGCGATAGTTTATTCTTAATCATATAAAATGAGATTTTCGTTATTATCTA
3720
3661 TATCGAAAGCAGCGCTATCAAAATAAGAAATTAGTATTTTACTCTAAAGCAATAATAGAT
3720
3721 CTGATAACCCCTTCAGATTATTCAACTCTACAGCCCTTTAACGCGAAAAAGCTTTCCTTTAT
3780
3721 GACTATTGGGAAGTCTAAAGTTGAGATGTCGGAAATTGCGCTTTTTCGAAAGGAAATA
3780
3781 ACCCAACCCATGCCGGGTTTTACTGGAGTGAAACCAGAAATACATAAACGGCAAAGGATGGC
3840
3781 TGGGTTGGGTACGGCCCAAATGACCTCACTTGCTCTTATGTATTTGCCCCTTTCCTACCG
3840
3841 AACGCTTCCGTTGCGGTTGCCGATCAGGCAAGGCGTATTTTTGAGGTGACGGTTAAACT
3900
3841 TTGCGAAGGCAACGCCAACGGCTAGTCCGTTCCGCATAAAATACTCCACTGCCAATTTGA
3900

Figure 12L

3901 TCCCGATCTCATTACTAAGAGCCACCTGCCATTAGATGATAGTATTTCAGTATGGCTGGA 3960
-----+-----+-----+-----+-----+-----+-----+
AGGGCTAGAGTAATGATTCTCGGTGGACGGTAATCTACTATCATAAGCTCATAACCGACCT
3961 TCAAAACAACCACTTATTGCCGTTTTTCATACATCCCGGCAAAAATACGTACACAGTTAG 4020
-----+-----+-----+-----+-----+-----+-----+
AGTTTGTGTTGGAATAACGGCAAAAGTATGTAGGGCCGTTTTTTTATGCAATGTGTCAATC
4021 AAAATGTAACGCTGCATGATGGATGCAGCAAAATCCCGGATTTCTGATATTACGCACAA 4080
-----+-----+-----+-----+-----+-----+-----+
TTTTACATTGCGACGTACTACCTACCGTCGTTTAAGGGCCCTAAAGACTATAAATGCCGTGTT
4081 CCTTGCAATGGCCCCGGATGGAGTCTGGTTACGCTGTACCCCATACGGTAATCTACATAATC 4140
-----+-----+-----+-----+-----+-----+-----+
GGAACGTACCGGGCCCTACCTCAGACCAATGCCACATGGGTATGCCATTAGATGTATTA
4141 GCATCTTAAAAAATTATCCTTCAACAAATCCCCCTTTACATTAAACAGCATTTGGTGTGATGA 4200
-----+-----+-----+-----+-----+-----+-----+
CGTAGAATTTTAAATAGGAAGTTGTTTAGGGGAAATGTAATTGTCGTAACCCAACTACT
4201 CGTCGGCTTTTTCCTGGTTACTACATCGCTCACTGGCCAAACCGTTATGGCGTTTGTGCG 4260
-----+-----+-----+-----+-----+-----+-----+
GCAGCCGAAACGACCAATGATGTAGCGAGTGACCGGTTTGGCAATACCGCAAAACAGC

Continuation of 09/201,945
PMS 10/COX(3)

Figure 12M

4261 ATGTCATTAAACCGCAACTGCACCGCTGAGCACACACGTTTACCAGCACAAACGACTGG 4320
-----+-----+-----+-----+-----+-----+-----+
TACAGTAATTATTTGGCGTTGACGTGGCGACTCGTGTGCAAAATGGTCGTGTTGCTGACC
4321 ATGAATTAGATAGTATTCGGTGCTTTTAAACCAACTGCTTGATACTCTACAAGTCCAAT 4380
-----+-----+-----+-----+-----+-----+-----+
TACTTAATCTATCATAACGGCCACGAAAATTGGTTGACGAACCTATGAGATGTTCAAGGTTA
4381 ACGACAAATCTGGAAAACAAAGTCGCAGACGCCAGCGCTAAATGAAGCAAAAACG 4440
-----+-----+-----+-----+-----+-----+-----+
TGCTGTTAGACCTTTTGTTCAGCGCTCGCGTGGTCCGCGATTACTTCGTTTTTTTGC
4441 CGCTGAGCNAGCTAACAAACGTAAAGCATTCATCTTACGGTAATAAGTCATGAGTTACG 4500
-----+-----+-----+-----+-----+-----+-----+
GCGACTCGNTCGATTGTTTGCAATTTTCGTAAGTAGAATGCCATTATTCAGTACTCAATGC
4501 TACTCCGATGAATGGCGTACTCGGTGCAATTGAATTATTACAAACCCCTTTAAACAT 4560
-----+-----+-----+-----+-----+-----+-----+
ATGAGGCTACTTACCGCATGAGCCACGTTAACTTAATAATGTTGGTGGGAAATTTGTA
4561 AGAGCAACAAGGATTAGCTGATACCGCCAGAAATTGTACACTGTCTTTGTAGCTATTAT 4620
-----+-----+-----+-----+-----+-----+-----+
TCTCGTTGTTCCCTAATCGACTATGGCGGTCTTTAAACATGTGACAGAAACAATCGATAATA

4621	TAATAATCTGCTGGATTTTTCACGCATCGAGTCTGGTCATTTTCACATTACATATGGAAGA -----+-----+-----+-----+-----+-----+-----+ ATTATTAGACGACCTAAAAAGTGCCTAGCTCAGACCAGTAAAGTGTAATGTATACCTTCT	4680
4681	AACAGCGTTACTGCCGTTACTGGACCAGGCAATGCAAAACCATCCAGGGCCAGCGCNA -----+-----+-----+-----+-----+-----+-----+ TTGTCCGAATGACGGCAATGACCTGGTCCGTTACGTTTGGTAGGTCCCCGGTCGCCGNTTT	4740
4741	GCAAAAACTGTCAATTACGTACTTTTGTCCGTCAACATGTCCCTCTCTATTTTCATACCG -----+-----+-----+-----+-----+-----+-----+ CGTTTTTTGACAGTAATGCATGAAAAACAGCCAGTTGTACAGGGAGAGATAAAAGTATGCG	4800
4801	ACAGTATCCGTTTACNNCAAATTTTGGTTAATTACTCGGAAACGCGTAAATTTACCG -----+-----+-----+-----+-----+-----+-----+ TGTCATAGGCAAATGNNGTTTAAACCAATTAAATGAGCCCTTGCGCCATTTTAAATGCG	4860
4861	AAACCGGAGGATACGTCTGACGGTCAAGCGTCAATGAGGAACAATTAATTTCTGGTTAG -----+-----+-----+-----+-----+-----+-----+ TTTGGCCCTCCTATGCAGACTGCCAGTTTCGCAGTACTCCTTGTTAATTATAAGACCAATC	4920
4921	CGATAGCCGGTAAAGGGATTGAAATACAGCAGCAGTCTCAAATCTTTACTGCTTTTATCA -----+-----+-----+-----+-----+-----+-----+ GCTATCGCCATTTCCTAACTTTATGTGTCGTCAGAGTTTAGAATGACGAAAAAATAGT	4980

4981
AGCAGACACAAAATTCGCAAGGATACAGGAAATGGACATGACATATGCGTCAAGCCCTGCGCTAA
-----+-----+-----+-----+-----+-----+-----+-----+
TCGTCTGTGTTTAAGCGTTCATGTCCCTTAACCTGACTGATAACGCAGTTTCGGACCGGATT
5040
AATGATGGCGGTAATCTGACACTAAAGTGTCCTCCGGGTTGGAACCTGTGTCTCGCT
5100
TACTACCCGCCATTAGACTGTGATTTTTCACAGGGGCCCCAACCTTGGACACAGAGCGA
5160
AGTATTACCCCTTACAAGAATACCAGCCGCTCAACCAATTAAAGGACGCTGTCAAGNNNC
-----+-----+-----+-----+-----+-----+-----+-----+
TCATAATGGGAATGTTCTATGTCGGCGAGTTGGTTAATTTCCCTGCGACAGTCNNNG
5220
CGTTCTGCCCTGCATCGGCAACTGGCTTGCTGGGGAATACGCGGTGAACCAACCCACCAGC
-----+-----+-----+-----+-----+-----+-----+-----+
GCAAGACGGACGTAGCCGTTGACCGAACGACCCCTTATGCGCCACTTGGTGGGTGTCG
5280
AAAATGCGCTTCTCAANNCNAGAGCTTTTGTATTTCTCCGGAAACTCTACGACCTGGCG
-----+-----+-----+-----+-----+-----+-----+-----+
TTTTACGCGAAGAGTTNNGNTCTCGAAAACATAAAGAGGCCCTTTTGAGATGCTGGACCGC
5340
CAACAGTTAATATTGTGTACACCAAAATATGCCAGTAATAATAATTGTTACCACCTGG
-----+-----+-----+-----+-----+-----+-----+-----+
GTTGTCAATTATAACACATGTGTTTATACGGTCATTATTATTATTAACAATGTTGGGACC

Figure 12P

5341 CAGTTGCAGATTCTTTTGGTTGATGATGCCGATATTAATCGGGATATCATCGGCAAAATG 5400
-----+-----+-----+-----+-----+-----+
GTCACGCTAAGAAAAACCAACTACTACGGCTATAATTAGCCCTATAGTAGCCGTTTAC
5401 CTTGTCAGCCTGGCCAAACACGTCACATAATTGCCGCCAGTAGTAACGAGGCTCTGACTTTA 5460
-----+-----+-----+-----+-----+-----+
GAACAGTCGACCCGGTTGTGCAGTGATAACGGCGGTCAATCATTTGCTCCGAGACTGAAAT
5461 TCACAAACAGCAGCGATTTCGATTTAGTACTGATTGACATTAGAATGCCAGAAATAGATGGT 5520
-----+-----+-----+-----+-----+-----+
AGTGTTCGTCGCTAAGCTAAATCATGACTAAGTAAATCTTACGGTCTTTATCTACCA
5521 ATTGAATGTGTACGATTATGGCATGATGAGCCGAATAATTAGATCCTGACTGCATGTTT 5580
-----+-----+-----+-----+-----+-----+
TAACTTACACATGCTAATAACCGTACTACTCGGCTTATTAAATCTAGGACTGACGTACAAA
5581 GTGGCACTATCCGCTAGCGTASCVNMAGAWRWTMTTCRTYGTDDAAAARWDRKDHWT 5640
-----+-----+-----+-----+-----+-----+
CACCGTGATAGCGGATCGCATSGBNKICTWYWAKWAGYARCAHHTTTTWTWYHCYMHWA
5641 CATHAYANNTTACAAAACCAGTGACATTGGCTACCTTAGCTCGCTACATCAGTATTGCCG 5700
-----+-----+-----+-----+-----+-----+
GTADTRTNNAATGTTTGGTCACTGTAAACCGATGGAATCGAGCGATGTAGTCATAACGGC
5701 CAGAAATACCAACTTTTACGAAATATAGAGCTACAGGAGCAGGATCC 5746
-----+-----+-----+-----+-----+-----+
GTCTTATGTTGAAAAATGCTTTTATATCTCGATGTCTCCTCGTCCCTAGG

Continuation of 09/201, 945
R0015101C02(13)